LAMN PROPERTY FOR THE DRIFT PARAMETER OF TIME INHOMOGENEOUS DIFFUSIONS WITH DISCRETE OBSERVATIONS

HOANG-LONG NGO AND NGOC KHUE TRAN

Abstract. We consider a multidimensional inhomogeneous diffusion whose drift coefficient depends on a multidimensional unknown parameter. Under some appropriate assumptions on the coefficients, we prove the local asymptotic mixed normality property for the drift parameter from high frequency observations when the length of the observation window tends to infinity. To obtain the result, we use the Malliavin calculus techniques and the Girsanov change of measures. Our approach is applicable for both ergodic and non-ergodic diffusions.

1. Introduction

We consider on a complete probability space $(\Omega, \mathcal{F}, \mathrm{P})$ a d-dimensional process $X^\theta = (X_t^\theta)_{t\geq 0}$ solution to the following inhomogeneous stochastic differential equation (SDE)

$$dX_t^\theta = b(\theta, t, X_t^\theta)dt + \sigma(t, X_t^\theta)dB_t, \tag{1.1}$$

where $X_0^\theta = x_0 \in \mathbb{R}^d$, and $B = (B_t)_{t\geq 0}$ is a d-dimensional Brownian motion. The unknown parameter $\theta = (\theta_1, \ldots, \theta_m)$ belongs to Θ, a compact subset of $\mathbb{R}^m$, for some integer $m \geq 1$.

Given $n \geq 1$, we consider a discrete observation scheme at deterministic and equidistant times $t_k = k\Delta_n$, $k \in \{0, \ldots, n\}$ of the process X^θ solution to (1.1), which is denoted by $X^{n,\theta} = (X_{t_0}^\theta, X_{t_1}^\theta, \ldots, X_{t_n}^\theta)$. We assume that the high-frequency and infinite horizon conditions hold. That is, $\Delta_n \to 0$ and $n\Delta_n \to \infty$ as $n \to \infty$. Let P_n^θ denote the probability law of the random vector $X^{n,\theta}$.

We say that the local asymptotic mixed normality (LAMN) property holds at $\theta^0 \in \Theta$ with asymptotic random Fisher information matrix $\Gamma(\theta^0)$ and rate of convergence $\varphi_{n\Delta_n}(\theta^0)$ if for any $u \in \mathbb{R}^m$, as $n \to \infty$,

$$\log \frac{d\mathrm{P}_n^{\theta^0+\varphi_{n\Delta_n}(\theta^0)u}}{d\mathrm{P}_n^{\theta^0}}\left(X^{n,\theta^0}\right) \overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})}{\longrightarrow} u^*\Gamma(\theta^0)^{1/2}\mathcal{N}(0, I_m) - \frac{1}{2}u^*\Gamma(\theta^0)u,$$

where $\mathcal{N}(0, I_m)$ is a centered $\mathbb{R}^m$-valued Gaussian random variable independent of $\Gamma(\theta^0)$ with identity covariance matrix I_m. Here, $\Gamma(\theta^0)$ is a symmetric positive definite random matrix in

Date: January 27, 2019.

2010 *Mathematics Subject Classification.* 60H07; 60J60; 62F12; 62M05.

Key words and phrases. asymptotic efficiency, inhomogeneous diffusion process, local asymptotic mixed normality property, Malliavin calculus, non-ergodic diffusion.

Ngoc Khue Tran wishes to thank the Vietnam Institute for Advanced Study in Mathematics (VIASM) for the intellectual environment and the hospitality where this work was done during his stay. This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 101.03-2017.316.

$\mathbb{R}^{m\times m}$, $\varphi_{n\Delta_n}(\theta^0)$ is a diagonal matrix in $\mathbb{R}^{m\times m}$ whose diagonal entries tend to zero as n goes to infinity, $\overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})}{\longrightarrow}$ denotes the convergence in $\widehat{\mathrm{P}}^{\theta^0}$-law which will be specified later on, and * denotes the transpose. If $\Gamma(\theta^0)$ is non-random, we say that the local asymptotic normality (LAN) property holds at θ^0.

The LAMN property plays a fundamental role in the asymptotic theory of statistics. This property developed by Jeganathan [12] extends the notion of LAN property introduced by Le Cam [15] and Hájek [8] in the situations where the asymptotic Fisher information matrix is deterministic. These properties allow to give the notion of asymptotically efficient estimators in the sense of Hájek-Le Cam convolution theorem as well as the lower bounds for the variance of estimators (see Jeganathan [12]). More precisely, a sequence of estimators $(\widehat{\theta}_n)_{n\geq 1}$ of the parameter θ^0 is called regular at θ^0 if for any $u\in\mathbb{R}^m$, as $n\to\infty$,

$$\varphi^{-1}_{n\Delta_n}(\theta^0)\left(\widehat{\theta}_n-\left(\theta^0+\varphi_{n\Delta_n}(\theta^0)u\right)\right)\overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0+\varphi_{n\Delta_n}(\theta^0)u})}{\longrightarrow}V(\theta^0),$$

for some $\mathbb{R}^m$-valued random variable $V(\theta^0)$, independent of u, where $\varphi^{-1}_{n\Delta_n}(\theta^0)$ denotes the inverse matrix of $\varphi_{n\Delta_n}(\theta^0)$. Note that taking $u=0$, this implies that as $n\to\infty$,

$$\varphi^{-1}_{n\Delta_n}(\theta^0)\left(\widehat{\theta}_n-\theta^0\right)\overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})}{\longrightarrow}V(\theta^0).$$

Suppose that the LAMN property holds at point θ^0. Let $(\widehat{\theta}_n)_{n\geq 1}$ be a regular sequence of estimators of the parameter θ^0. Then the law of $V(\theta^0)$ conditionally on $\Gamma(\theta^0)$ is a convolution between the Gaussian law $\mathcal{N}(0,\Gamma(\theta^0)^{-1})$ and some other law $G_{\Gamma(\theta^0)}$ on $\mathbb{R}^m$, that is,

$$\mathcal{L}\left(V(\theta^0)|\Gamma(\theta^0)\right)=\mathcal{N}\left(0,\Gamma(\theta^0)^{-1}\right)\star G_{\Gamma(\theta^0)}.$$

Hence, the random variable $V(\theta^0)$ can be written as a sum of two independent random variables

$$V(\theta^0)\overset{\text{law}}{=}\Gamma(\theta^0)^{-1/2}\mathcal{N}(0,I_m)+R,$$

where R is a random variable with distribution $G_{\Gamma(\theta^0)}$, independent of $\mathcal{N}(0,I_m)$ (see [12, Corollary 1]). This implies that as $n\to\infty$,

$$\varphi^{-1}_{n\Delta_n}(\theta^0)\left(\widehat{\theta}_n-\theta^0\right)\overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})}{\longrightarrow}\Gamma(\theta^0)^{-1/2}\mathcal{N}(0,I_m)+R.$$

This conditional convolution theorem suggests the notion of asymptotically efficient estimators in terms of minimal asymptotic variance when $R=0$. That is, assume that the LAMN property holds at point θ^0, a sequence of estimators $(\widehat{\theta}_n)_{n\geq 1}$ of the parameter θ^0 is called asymptotically efficient at θ^0 in the sense of Hájek-Le Cam convolution theorem if as $n\to\infty$,

$$\varphi^{-1}_{n\Delta_n}(\theta^0)\left(\widehat{\theta}_n-\theta^0\right)\overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})}{\longrightarrow}\Gamma(\theta^0)^{-1/2}\mathcal{N}(0,I_m),$$

where $\Gamma(\theta^0)$ and $\mathcal{N}(0,I_m)$ are independent. We refer the reader to Subsection 7.1 of Höpfner [9] or Le Cam and Lo Yang [16] for further details.

On the basis of continuous observations with increasing observation window, the LAMN property was established by Luschgy in [17] for semimartingale, by Kutoyants in [13] for null-recurrent process (see [13, Remark 3.42]), and for Ornstein-Uhlenbeck process (see [13, Remark 3.47]), and by Bishwal in [4, Chapter 4] for inhomogeneous diffusions. Let us mention here that the asymptotic likelihood theory for multidimensional inhomogeneous diffusion processes (1.1) whose drift coefficient depends linearly on the parameter can be found in Section 5 of [1, Chapter 9], which includes the case of homogeneous ergodic diffusions. Besides, the asymptotic properties of maximum likelihood estimator and Bayes estimator for the nonlinear drift parameter of one-dimensional inhomogeneous diffusion were also studied in [4, Chapter 4] and [18]. In [6], Gobet proved the LAMN property for elliptic diffusion based on discrete observations on a fixed time interval. Later on, from discrete observations with increasing observation window $n\Delta_n$, Gobet in [7] obtained the LAN property for homogeneous ergodic diffusions using Malliavin calculus, and Shimizu in [20] showed the LAMN property for a particular case of non-recurrent Ornstein-Uhlenbeck process using the explicit expression of the transition density. Recall also that results on parameter estimation for discretely observed non-ergodic diffusions can be found in Jacod [10] where the rate is $(\sqrt{n\Delta_n}, \sqrt{n})$ for the drift and diffusion parameters, respectively, and in Shimizu [21] where the rate varies depending on the observed Fisher information. Indeed, in [10], the author constructed estimators from a moment type contrast function for the drift and diffusion parameters of multidimensional homogeneous and non-ergodic diffusions and established the consistency of the estimators in the sense of tightness under some suitable smoothness and identifiability conditions. These estimators converge at rate $\sqrt{n\Delta_n}$ for the drift parameter and at rate $\sqrt{n}$ for the diffusion parameter. In [21], the author constructed M-estimators from a quadratic-type contrast function for the drift and diffusion parameters of one-dimensional homogeneous diffusions without ergodicity assumption and established the consistency of the M-estimators in the sense of tightness. These M-estimators converge with a variety of rates of convergence for the drift and diffusion parameters. However, the validity of the LAMN property on the basis of discrete observations of solution to a general inhomogeneous and non-ergodic SDE when the length of the observation window tends to infinity has not been investigated yet.

In this paper, we prove the LAMN property for a general class of inhomogeneous diffusions observed at discrete time without assuming ergodicity. Unlike the Ornstein-Uhlenbeck process, the transition density of the solution to the general equation (1.1) is not explicit. Therefore, we use the Malliavin calculus approach initiated by Gobet [6] to derive an explicit expression for the logarithm derivative of the transition density w.r.t. the parameter (see Lemma 3.3). With the help of this explicit expression, we derive an appropriate expansion of the log-likelihood ratio (see Lemma 4.1). In order to treat the main contributions, we need to use the asymptotic behavior of the observed Fisher information process based on the continuous observation (see condition **(A4)** below) together with the multivariate central limit theorem for continuous local martingales (see Lemma 4.2). As will be seen in Subsection 4.3, with the help of two conditions **(A5)**-**(A6)**, the negligible contribution of the expansion is shown by using two technical Lemmas 3.6 and 3.7 which are respectively related to the Girsanov change of measures and the deviation of Girsanov change of measures when the drift parameter changes. This techniques is not the same as the one that Gobet used in [7]. Indeed, in [7] the author used a change of transition densities, the upper and lower bounds of Gaussian type of the transition densities together with the ergodic property. In our situation, the ergodic assumption is not required, which makes impossible to implement the argument in Gobet [7].

The paper is organized as follows. In Section 2, we formulate the assumptions on equation (1.1) and state our main result in Theorem 2.1. Section 3 presents preliminary results needed for the proof of the main result, which concern the explicit expression for the logarithm derivative of the transition density w.r.t. the parameter and the Girsanov change of measures. The proofs of these technical results are postponed to Appendix in Section 6 in order to maintain the flow of the exposition. We prove our main result in Section 4, which follows the aforementioned strategy. Several illustrated examples will be also given in Section 5 which discusses homogeneous ergodic diffusion processes, homogeneous Ornstein-Uhlenbeck process, two-dimensional Gaussian diffusion process, null-recurrent diffusion process, exponential growth process, inhomogeneous Ornstein-Uhlenbeck process and a special inhomogeneous diffusion process.

2. Assumptions and main result

Let $\{\widehat{\mathcal{F}}_t\}_{t\geq 0}$ denote the natural filtration generated by B. We always suppose that the coefficients $b=(b_1,\ldots,b_d):\Theta\times\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma:\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^d$ are measurable functions satisfying the Lipschitz continuity and linear growth condition **(A1)** below under which equation (1.1) has a unique $\{\widehat{\mathcal{F}}_t\}_{t\geq 0}$-adapted solution X^θ possessing the strong Markov property. We denote by $\widehat{\mathrm{P}}^\theta$ the probability measure induced by the process X^θ on the canonical space $(C(\mathbb{R}_+,\mathbb{R}^d),\mathcal{B}(C(\mathbb{R}_+,\mathbb{R}^d)))$ endowed with the natural filtration $\{\widehat{\mathcal{F}}_t\}_{t\geq 0}$. Here $C(\mathbb{R}_+,\mathbb{R}^d)$ denotes the set of $\mathbb{R}^d$-valued continuous functions defined on $\mathbb{R}_+$, and $\mathcal{B}(C(\mathbb{R}_+,\mathbb{R}^d))$ is its Borel σ-algebra. We denote by $\widehat{\mathrm{E}}^\theta$ the expectation with respect to (w.r.t.) $\widehat{\mathrm{P}}^\theta$. Let $\overset{\widehat{\mathrm{P}}^\theta}{\longrightarrow}$, $\overset{\mathcal{L}(\widehat{\mathrm{P}}^\theta)}{\longrightarrow}$, $\widehat{\mathrm{P}}^\theta$-a.s., $\overset{\mathrm{P}}{\longrightarrow}$, and $\overset{\mathcal{L}(\mathrm{P})}{\longrightarrow}$ denote the convergence in $\widehat{\mathrm{P}}^\theta$-probability, in $\widehat{\mathrm{P}}^\theta$-law, in $\widehat{\mathrm{P}}^\theta$-almost surely, in P-probability, and in P-law, respectively. For $x\in\mathbb{R}^d$, $|x|$ denotes the Euclidean norm. $|A|$ denotes the Frobenius norm of the square matrix A, and $\mathrm{tr}(A)$ denotes the trace.

We now recall some concepts on asymptotic statistical inference for the continuously observed parametric model. For details, we refer the reader to Barndorff-Nielsen and Sørensen [3]. For any $T\geq 0$ and $\theta\in\Theta$, we let $\widehat{\mathrm{P}}_T^\theta$ denote the probability measure generated by the process $X^{T,\theta}:=(X_t^\theta)_{t\in[0,T]}$ solving equation (1.1) under the parameter θ on the measurable space $(C([0,T],\mathbb{R}^d),\mathcal{B}(C([0,T],\mathbb{R}^d)))$. Here $C([0,T],\mathbb{R}^d)$ denotes the set of $\mathbb{R}^d$-valued continuous functions defined on $[0,T]$, and $\mathcal{B}(C([0,T],\mathbb{R}^d))$ is its Borel σ-algebra. Therefore, $\widehat{\mathrm{P}}_T^\theta$ is the restriction of $\widehat{\mathrm{P}}^\theta$ to $\widehat{\mathcal{F}}_T$. We define the log-likelihood function of the family of probability measures $(\widehat{\mathrm{P}}_T^\theta)_{\theta\in\Theta}$ as

$$\ell_T(\theta)=\log\frac{d\widehat{\mathrm{P}}_T^\theta}{d\widehat{\widehat{\mathrm{P}}}_T},$$

where $\widehat{\widehat{\mathrm{P}}}_T$ is a probability measure on $(C([0,T],\mathbb{R}^d),\mathcal{B}(C([0,T],\mathbb{R}^d)))$ which is supposed to satisfy that $\widehat{\mathrm{P}}_T^\theta$ is absolutely continuous w.r.t. $\widehat{\widehat{\mathrm{P}}}_T$, for all $T\geq 0$ and $\theta\in\Theta$. In fact, by [13, Theorem 1.12], for all $\theta,\theta^1\in\Theta$, the probability measures $\widehat{\mathrm{P}}_T^\theta$ and $\widehat{\mathrm{P}}_T^{\theta^1}$ are absolutely continuous w.r.t. each other and its Radon-Nikodym derivative is given by

$$\frac{d\widehat{\mathrm{P}}_T^\theta}{d\widehat{\mathrm{P}}_T^{\theta^1}}\left((X_t^{\theta^1})_{t\in[0,T]}\right)=\exp\left\{\int_0^T\sigma^{-1}(t,X_t^{\theta^1})\left(b(\theta,t,X_t^{\theta^1})-b(\theta^1,t,X_t^{\theta^1})\right)\cdot dB_t\right.$$
$$\left.-\frac{1}{2}\int_0^T\left|\sigma^{-1}(t,X_t^{\theta^1})\left(b(\theta,t,X_t^{\theta^1})-b(\theta^1,t,X_t^{\theta^1})\right)\right|^2dt\right\}.$$

By Girsanov's theorem, the process $\widehat{B} = (\widehat{B}_t)_{t\in[0,T]}$ is a Brownian motion under $\widehat{\mathrm{P}}^\theta$, where for any $t \in [0,T]$,

$$\widehat{B}_t := B_t - \int_0^t \sigma^{-1}(s, X_s^{\theta^1}) \left(b(\theta, s, X_s^{\theta^1}) - b(\theta^1, s, X_s^{\theta^1})\right) ds.$$

Therefore, the log-likelihood function is given by

$$\begin{aligned}\ell_T(\theta) = \log \frac{d\widehat{\mathrm{P}}_T^\theta}{d\widehat{\mathrm{P}}_T^{\theta^1}} \left((X_t^{\theta^1})_{t\in[0,T]}\right) &= \int_0^T \sigma^{-1}(t, X_t^{\theta^1}) \left(b(\theta, t, X_t^{\theta^1}) - b(\theta^1, t, X_t^{\theta^1})\right) \cdot dB_t \\ &- \frac{1}{2}\int_0^T \left|\sigma^{-1}(t, X_t^{\theta^1}) \left(b(\theta, t, X_t^{\theta^1}) - b(\theta^1, t, X_t^{\theta^1})\right)\right|^2 dt,\end{aligned}$$

where $\widehat{\mathrm{P}}_T^{\theta^1}$ is considered as the dominating probability measure $\widehat{\widehat{\mathrm{P}}}_T$ of the family of probability measures $(\widehat{\mathrm{P}}_T^\theta)_{\theta\in\Theta}$. The score vector which is defined as the vector of first derivatives of the log-likelihood function is given by the gradient

$$\nabla_\theta \ell_T(\theta) = \int_0^T \sigma^{-1}(t, X_t^{\theta^1}) \nabla_\theta b(\theta, t, X_t^{\theta^1}) \cdot \left(dB_t - \sigma^{-1}(t, X_t^{\theta^1}) \left(b(\theta, t, X_t^{\theta^1}) - b(\theta^1, t, X_t^{\theta^1})\right)\right).$$

Hence, under $\widehat{\mathrm{P}}^\theta$, the score vector is rewritten as

$$\nabla_\theta \ell_T(\theta) = \int_0^T \sigma^{-1}(t, X_t^\theta) \nabla_\theta b(\theta, t, X_t^\theta) d\widehat{B}_t,$$

which is a martingale w.r.t. the filtration $\{\widehat{\mathcal{F}}_t\}_{t\in[0,T]}$. The quadratic variation of the score vector martingale is given by

$$[\nabla_\theta \ell(\theta)]_T = \int_0^T (\nabla_\theta b(\theta, t, X_t^\theta))^* (\sigma^{-1}(t, X_t^\theta))^* \sigma^{-1}(t, X_t^\theta) \nabla_\theta b(\theta, t, X_t^\theta) dt,$$

which can be interpreted as the observed Fisher information process at θ based on the continuous observation $(X_t^\theta)_{t\in[0,T]}$.

We impose the following assumptions on equation (1.1).

(A1) For any $\theta \in \Theta$, there exist a constant $L > 0$ such that for all $x, y \in \mathbb{R}^d$ and $t \geq 0$,

$$\begin{aligned}&|b(\theta, t, x) - b(\theta, t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq L|x - y|, \\ &|b(\theta, t, x)| + |\sigma(t, x)| \leq L(1 + |x|).\end{aligned}$$

Moreover, the Lipschitz constant L is uniformly bounded on Θ.

(A2) The diffusion matrix σ is symmetric, positive and satisfies an uniform ellipticity condition, that is, there exists a constant $c \geq 1$ such that for all $x, \xi \in \mathbb{R}^d$ and $t \geq 0$,

$$\frac{1}{c}|\xi|^2 \leq |\sigma(t, x)\xi|^2 \leq c|\xi|^2.$$

(A3) The functions b and σ are of class C^1 w.r.t. θ, t and x. Each partial derivative $\partial_{\theta_i} b$, $\partial_{x_i} b$ and $\partial_{x_i}\sigma$ is of class C^1 w.r.t. x, and $\partial_{\theta_i} b$ is of class C^1 w.r.t. t. Moreover, for any $(\theta, \theta_1, \theta_2, x, y) \in \Theta^3 \times (\mathbb{R}^d)^2$ and $t \geq 0$, there exist positive constants C, γ, independent of $(\theta, \theta_1, \theta_2, x, y)$ such that

(a) $|g(\cdot,t,x)| \leq C$ for $g(\cdot,t,x) = \partial_{x_i} b(\theta,t,x), \partial_{x_i}\sigma(t,x), \partial_t \sigma(t,x), \partial^2_{\theta_i x_j} b(\theta,t,x), \partial^2_{x_i x_j} b(\theta,t,x), \partial^2_{x_i x_j}\sigma(t,x)$;

(b) $|h(\cdot,t,x)| \leq C\,(1+|x|)$ for $h(\cdot,x) = \partial_{\theta_i} b(\theta,t,x), \partial_t b(\theta,t,x), \partial^2_{\theta_i t} b(\theta,t,x)$;

(c) $|\partial_{\theta_i} b(\theta_1,t,x) - \partial_{\theta_i} b(\theta_2,t,x)| \leq C|\theta_1-\theta_2|^{\gamma}\,(1+|x|)$.

(A4) For any $\theta \in \Theta$, there exist a $m\times m$ non-random diagonal matrix $\varphi_T(\theta) = \mathrm{diag}(\varphi^1_T(\theta),\ldots,\varphi^m_T(\theta))$ whose diagonal entries $\varphi^1_T(\theta),\ldots,\varphi^m_T(\theta)$ are strictly positive, decreasing w.r.t. T and tend to zero as $T\to\infty$, and a $m\times m$ symmetric positive definite random matrix $\Gamma(\theta)$ such that the observed Fisher information process at θ based on the continuous observation $(X^\theta_t)_{t\in[0,T]}$ converges to $\Gamma(\theta)$ at rate $\varphi_T(\theta)$ in $\widehat{\mathrm{P}}^\theta$-probability as $T\to\infty$. That is, as $T\to\infty$,

$$\varphi_T(\theta)\int_0^T (\nabla_\theta b(\theta,t,X^\theta_t))^*(\sigma^{-1}(t,X^\theta_t))^*\sigma^{-1}(t,X^\theta_t)\nabla_\theta b(\theta,t,X^\theta_t)dt\;\varphi_T(\theta) \overset{\widehat{\mathrm{P}}^\theta}{\longrightarrow} \Gamma(\theta).$$

(A5) For any $\theta\in\Theta$ and $i,j\in\{1,\ldots,m\}$,

$$\widehat{\mathrm{E}}^\theta\left[\sup_{t\geq 0}|\varphi^i_t(\theta)X^\theta_t|\right] + \widehat{\mathrm{E}}^\theta\left[\sup_{t\geq 0}\varphi^i_t(\theta)\varphi^j_t(\theta)|X^\theta_t|^2\right] + \widehat{\mathrm{E}}^\theta\left[\sup_{t\geq 0}(\varphi^i_t(\theta))^2\varphi^j_t(\theta)|X^\theta_t|^3\right] + \widehat{\mathrm{E}}^\theta\left[\sup_{t\geq 0}|\varphi^i_t(\theta)X^\theta_t|^4\right] < \infty.$$

(A6) For any $\theta\in\Theta$ and $i\in\{1,\ldots,m\}$, as $n\to\infty$,

$$n\Delta_n^2 \to 0,\; n\Delta_n^{\frac{3}{2}}\varphi^i_{n\Delta_n}(\theta)\to 0,\; n\Delta_n^{\frac{5}{2}}(\varphi^i_{n\Delta_n}(\theta))^{-1}\to 0,\; n\Delta_n^4(\varphi^i_{n\Delta_n}(\theta))^{-2}\to 0.$$

To be able to apply the Malliavin calculus, the uniform ellipticity condition **(A2)** and regularity condition **(A3)** on the coefficients are required. Condition **(A4)** is given in order to ensure asymptotic result for the score vector, which will be seen in Subsection 4.2. Let us recall that condition **(A4)** is similar to general condition (3.3) of Barndorff-Nielsen and Sørensen [3] which is given for general asymptotic likelihood theory for stochastic processes. This condition **(A4)** is also similar to condition (2.12) of Luschgy [17] which is established for semimartingales. It is worth noticing that from page 155 of Luschgy [17], the chosen rate $\varphi_T(\theta)$ in condition **(A4)** is naturally concerned with the expected Fisher information at θ and T based on the continuous observation $(X^\theta_t)_{t\in[0,T]}$, which is defined by

$$\widehat{\mathrm{E}}^\theta\left[[\nabla_\theta \ell(\theta)]_T\right] = \widehat{\mathrm{E}}^\theta\left[\int_0^T (\nabla_\theta b(\theta,t,X^\theta_t))^*(\sigma^{-1}(t,X^\theta_t))^*\sigma^{-1}(t,X^\theta_t)\nabla_\theta b(\theta,t,X^\theta_t)dt\right].$$

Let us mention that as will be seen in Subsection 4.3, conditions **(A5)** and **(A6)** are due to our techniques developed in this paper, which are used to show the negligible contributions for negligible terms in the expansion of the log-likelihood ratio. They are similar to condition (48) on page 572 and conditions (13), (14), (16) on page 554 and 555 of [21] which are used to prove the consistency of the aforementioned M-estimators.

Now, for fixed $\theta^0\in\Theta$, we consider a discrete observation $X^{n,\theta^0} = (X^{\theta^0}_{t_0}, X^{\theta^0}_{t_1},\ldots,X^{\theta^0}_{t_n})$ of the process X^{θ^0}. The main result of this paper is the following LAMN property.

Theorem 2.1. *Assume conditions* **(A1)**-**(A6)**. *Then, the LAMN property holds for the likelihood at* θ^0 *with rate of convergence* $\varphi_{n\Delta_n}(\theta^0) = \mathrm{diag}(\varphi^1_{n\Delta_n}(\theta^0),\ldots,\varphi^m_{n\Delta_n}(\theta^0))$ *and asymptotic* *random Fisher information matrix* $\Gamma(\theta^0)$*. That is, for all* $u \in \mathbb{R}^m$*, as* $n \to \infty$*,*

$$\log \frac{d\mathrm{P}_n^{\theta^0+\varphi_{n\Delta_n}(\theta^0)u}}{d\mathrm{P}_n^{\theta^0}}\left(X^{n,\theta^0}\right) \overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})}{\longrightarrow} u^*\Gamma(\theta^0)^{1/2}\mathcal{N}(0,I_m) - \frac{1}{2}u^*\Gamma(\theta^0)u,$$

where $\mathcal{N}(0,I_m)$ *is a centered* $\mathbb{R}^m$*-valued Gaussian random variable independent of* $\Gamma(\theta^0)$ *with identity covariance matrix* I_m*.*

Technical conditions **(A5)** and **(A6)** now can be simplified for some particular classes of diffusions, which depends on the homogeneity of the coefficients and the derivatives of drift coefficient. See Remark 4.8.

When equation (1.1) is time homogeneous, then condition **(A3)**(b) becomes $\partial_t b(\theta,x) = \partial^2_{\theta_i t} b(\theta,x) = 0$ and $|\partial_{\theta_i} b(\theta,x)| \leq C(1+|x|)$. In this case, conditions **(A5)**-**(A6)** are reformulated as follows

(A5') For any $\theta \in \Theta$ and $i,j \in \{1,\ldots,m\}$,

$$\widehat{\mathrm{E}}^{\theta}\left[\sup_{t\geq 0}|\varphi_t^i(\theta)X_t^{\theta}|\right] + \widehat{\mathrm{E}}^{\theta}\left[\sup_{t\geq 0}\varphi_t^i(\theta)\varphi_t^j(\theta)|X_t^{\theta}|^2\right] < \infty.$$

(A6') For any $\theta \in \Theta$ and $i \in \{1,\ldots,m\}$, as $n \to \infty$,

$$n\Delta_n^2 \to 0,\ n\Delta_n^{\frac{3}{2}}\varphi^i_{n\Delta_n}(\theta) \to 0.$$

Corollary 2.2. *Let equation* (1.1) *be homogeneous. Assume conditions* **(A1)**-**(A4)** *and* **(A5')**-**(A6')***. Then, the statement of Theorem 2.1 remains valid.*

Furthermore, when $|\partial_{\theta_i} b(\theta,x)| \leq C|x|$ for all $\theta \in \Theta$, $i \in \{1,\ldots,m\}$, $x \in \mathbb{R}^d$ and some constant $C > 0$, condition **(A6)** is reformulated as follows

(A6") For any $\theta \in \Theta$ and $i \in \{1,\ldots,m\}$, $n\Delta_n^{\frac{3}{2}}\varphi^i_{n\Delta_n}(\theta) \to 0$ as $n \to \infty$.

Corollary 2.3. *Let equation* (1.1) *be homogeneous. Assume that* $|\partial_{\theta_i} b(\theta,x)| \leq C|x|$ *for all* $\theta \in \Theta$*,* $i \in \{1,\ldots,m\}$*,* $x \in \mathbb{R}^d$ *and some constant* $C > 0$*, and assume conditions* **(A1)**-**(A4)** *and* **(A5')**-**(A6")***. Then, the statement of Theorem 2.1 remains valid.*

When equation (1.1) is homogeneous and $\partial_{\theta_i} b(\theta,x)$ is bounded, in this case, condition **(A5)** is not required and condition **(A6)** is reformulated as follows

(A6"') For any $\theta \in \Theta$ and $i,j \in \{1,\ldots,m\}$, $n\Delta_n^{\frac{3}{2}}\varphi^i_{n\Delta_n}(\theta)\varphi^j_{n\Delta_n}(\theta) \to 0$ as $n \to \infty$.

Corollary 2.4. *Let equation* (1.1) *be homogeneous. Assume that* $\partial_{\theta_i} b(\theta,x)$ *is bounded for all* $i \in \{1,\ldots,m\}$*, and assume conditions* **(A1)**-**(A4)** *and* **(A6"')***. Then, the statement of Theorem 2.1 remains valid.*

Remark 2.5. *Theorem 2.1 can be seen as an extension of* [7, Theorem 4.1] *when the unknown parameter appears only in the drift coefficient and when equation* (1.1) *is homogeneous and ergodic (see Subsection 5.1.1).*

As usual, constants will be denoted by C and they will always be independent of time and Δ_n but may depend on bounds for the set Θ. They may change of value from one line to the next.

3. Preliminaries

In this section, we introduce some preliminary results needed for the proof of Theorem 2.1. For this, we consider the canonical filtered probability spaces $(\widehat{\Omega}, \widehat{\mathcal{F}}, \{\widehat{\mathcal{F}}_t\}_{t\geq 0}, \widehat{\mathrm{P}})$ and $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \{\widetilde{\mathcal{F}}_t\}_{t\geq 0}, \widetilde{\mathrm{P}})$ associated respectively to each of two processes B and W, where $W = (W_t)_{t\geq 0}$ is a d-dimensional standard Brownian motion independent of B. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathrm{P})$ be the product filtered probability space of these two canonical spaces. That is, $\Omega = \widehat{\Omega} \times \widetilde{\Omega}$, $\mathcal{F} = \widehat{\mathcal{F}} \otimes \widetilde{\mathcal{F}}$, $\mathrm{P} = \widehat{\mathrm{P}} \otimes \widetilde{\mathrm{P}}$, $\mathcal{F}_t = \widehat{\mathcal{F}}_t \otimes \widetilde{\mathcal{F}}_t$, and $\mathrm{E} = \widehat{\mathrm{E}} \otimes \widetilde{\mathrm{E}}$, where E, $\widehat{\mathrm{E}}$, $\widetilde{\mathrm{E}}$ denote the expectation w.r.t. P, $\widehat{\mathrm{P}}$ and $\widetilde{\mathrm{P}}$, respectively.

To simplify the exposition, for $i \in \{1, \ldots, m\}$ we set

$$\begin{aligned}
&\theta^0 = (\theta^0_1, \ldots, \theta^0_m), u = (u_1, u_2, \ldots, u_m),\\
&\theta^{0+} := \theta^0 + \varphi_{n\Delta_n}(\theta^0)u = (\theta^0_1 + \varphi^1_{n\Delta_n}(\theta^0)u_1, \ldots, \theta^0_m + \varphi^m_{n\Delta_n}(\theta^0)u_m),\\
&\theta^{0+}_i := (\theta^0_1, \ldots, \theta^0_{i-1}, \theta^0_i + \varphi^i_{n\Delta_n}(\theta^0)u_i, \theta^0_{i+1} + \varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}, \ldots, \theta^0_m + \varphi^m_{n\Delta_n}(\theta^0)u_m),\\
&\theta^{0+}_i(\ell) := (\theta^0_1, \ldots, \theta^0_{i-1}, \theta^0_i + \ell\varphi^i_{n\Delta_n}(\theta^0)u_i, \theta^0_{i+1} + \varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}, \ldots, \theta^0_m + \varphi^m_{n\Delta_n}(\theta^0)u_m).
\end{aligned}$$

Under conditions **(A1)**, **(A2)** and **(A3)**(a), for any $t > s$ the law of X^θ_t conditioned on $X^\theta_s = x$ admits a positive transition density $p^\theta(s, t, x, y)$, which is differentiable w.r.t. θ. We denote by $p_n(\cdot; \theta)$ the density of the random vector $X^{n,\theta}$. To deal with the log-likelihood ratio in Theorem 2.1, we use the Markov property to rewrite the global likelihood function in terms of a product of transition densities and then apply a mean value theorem. Precisely,

$$\begin{aligned}
&\log \frac{d\mathrm{P}_n^{\theta^0+\varphi_{n\Delta_n}(\theta^0)u}}{d\mathrm{P}_n^{\theta^0}}\left(X^{n,\theta^0}\right) = \log \frac{p_n\left(X^{n,\theta^0}; \theta^0 + \varphi_{n\Delta_n}(\theta^0)u\right)}{p_n\left(X^{n,\theta^0}; \theta^0\right)} = \log \frac{p_n\left(X^{n,\theta^0}; \theta^{0+}\right)}{p_n\left(X^{n,\theta^0}; \theta^0\right)}\\
&= \sum_{k=0}^{n-1} \log \frac{p^{\theta^{0+}}}{p^{\theta^0}}\left(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}}\right) = \sum_{k=0}^{n-1} \log \frac{p^{\theta^{0+}_1}}{p^{\theta^0}}\left(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}}\right)\\
&= \sum_{k=0}^{n-1} \log \left(\frac{p^{\theta^{0+}_1}}{p^{\theta^{0+}_2}} \frac{p^{\theta^{0+}_2}}{p^{\theta^{0+}_3}} \cdots \frac{p^{\theta^{0+}_i}}{p^{\theta^{0+}_{i+1}}} \cdots \frac{p^{\theta^{0+}_{m-1}}}{p^{\theta^{0+}_m}} \frac{p^{\theta^{0+}_m}}{p^{\theta^0}}\right)\left(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}}\right)\\
&= \sum_{k=0}^{n-1} \log \frac{p^{\theta^{0+}_1}}{p^{\theta^{0+}_2}}\left(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}}\right) + \sum_{k=0}^{n-1} \log \frac{p^{\theta^{0+}_2}}{p^{\theta^{0+}_3}}\left(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}}\right)\\
&\qquad + \cdots + \sum_{k=0}^{n-1} \log \frac{p^{\theta^{0+}_i}}{p^{\theta^{0+}_{i+1}}}\left(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}}\right) + \cdots + \sum_{k=0}^{n-1} \log \frac{p^{\theta^{0+}_m}}{p^{\theta^0}}\left(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}}\right)\\
&= \sum_{k=0}^{n-1} \varphi^1_{n\Delta_n}(\theta^0)u_1 \int_0^1 \frac{\partial_{\theta_1} p^{\theta^{0+}_1(\ell)}}{p^{\theta^{0+}_1(\ell)}}\left(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}}\right) d\ell\\
&\qquad + \sum_{k=0}^{n-1} \varphi^2_{n\Delta_n}(\theta^0)u_2 \int_0^1 \frac{\partial_{\theta_2} p^{\theta^{0+}_2(\ell)}}{p^{\theta^{0+}_2(\ell)}}\left(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}}\right) d\ell\\
&\qquad + \cdots + \sum_{k=0}^{n-1} \varphi^i_{n\Delta_n}(\theta^0)u_i \int_0^1 \frac{\partial_{\theta_i} p^{\theta^{0+}_i(\ell)}}{p^{\theta^{0+}_i(\ell)}}\left(t_k, t_{k+1}, X^{\theta^0}_{t_k}, X^{\theta^0}_{t_{k+1}}\right) d\ell
\end{aligned}$$

$$+\cdots+\sum_{k=0}^{n-1}\varphi_{n\Delta_n}^{m}(\theta^0)u_m\int_0^1\frac{\partial_{\theta_m}p^{\theta_m^{0+}(\ell)}}{p^{\theta_m^{0+}(\ell)}}\left(t_k,t_{k+1},X_{t_k}^{\theta^0},X_{t_{k+1}}^{\theta^0}\right)d\ell. \tag{3.1}$$

We start as in Gobet [6] by applying the integration by parts formula of the Malliavin calculus on each interval $[t_k,t_{k+1}]$ to obtain an explicit expression for the logarithm derivative of the transition density w.r.t. the parameter. In order to avoid confusion with the observed process X^θ, we introduce an extra probabilistic representation of X^θ for which the Malliavin calculus will be applied. That is, we consider on the same probability space $(\Omega,\mathcal{F},\mathrm{P})$ the stochastic flow $Y^\theta(s,x)=(Y_t^\theta(s,x),t\geq s)$, $x\in\mathbb{R}^d$ on the time interval $[s,\infty)$ and with initial condition $Y_s^\theta(s,x)=x$ satisfying

$$Y_t^\theta(s,x)=x+\int_s^t b(\theta,u,Y_u^\theta(s,x))du+\int_s^t\sigma(u,Y_u^\theta(s,x))dW_u. \tag{3.2}$$

In particular, we write $Y_t^\theta\equiv Y_t^\theta(0,x_0)$, for all $t\geq 0$. That is,

$$Y_t^\theta=x_0+\int_0^t b(\theta,u,Y_u^\theta)du+\int_0^t\sigma(u,Y_u^\theta)dW_u. \tag{3.3}$$

We will apply the Malliavin calculus on the Wiener space induced by W. Let D and δ denote the Malliavin derivative and the Skorohod integral w.r.t. W on each interval $[t_k,t_{k+1}]$, respectively. We denote by $\mathbb{D}^{1,2}$ the space of random variables differentiable in the sense of Malliavin, and by Dom δ the domain of δ. We refer to Nualart [19] for a detailed exposition of the Malliavin calculus on the Wiener space. Recall that for a differentiable random variable $F\in\mathbb{D}^{1,2}$, its Malliavin derivative is denoted by $DF=(D^1F,\ldots,D^dF)$, where D^i is the Malliavin derivative in the ith direction W^i of the Brownian motion $W=(W^1,\ldots,W^d)$, for $i\in\{1,\ldots,d\}$. For a $\mathbb{R}^d$-valued process $U=(U^1,\ldots,U^d)\in\text{Dom }\delta$, the Skorohod integral of U is defined as $\delta(U)=\sum_{i=1}^d\delta^i(U^i)$, where δ^i denotes the Skorohod integral w.r.t. W^i.

For any $k\in\{0,\ldots,n-1\}$, under conditions **(A1)**, **(A2)** and **(A3)**(a)-(b), the process $(Y_t^\theta(t_k,x),t\in[t_k,t_{k+1}])$ is differentiable w.r.t. x and θ. We denote by $(\nabla_xY_t^\theta(t_k,x),t\in[t_k,t_{k+1}])$ the Jacobian matrix, and by $(\partial_{\theta_i}Y_t^\theta(t_k,x),t\in[t_k,t_{k+1}])$ the derivative w.r.t. θ_i for $i\in\{1,\ldots,m\}$ (see Kunita [14]). These processes are the solutions to the linear equations

$$\begin{aligned}\nabla_xY_t^\theta(t_k,x)&=\mathrm{I}_d+\int_{t_k}^t\nabla_xb(\theta,s,Y_s^\theta(t_k,x))\nabla_xY_s^\theta(t_k,x)ds \\ &\quad+\sum_{j=1}^d\int_{t_k}^t\nabla_x\sigma_j(s,Y_s^\theta(t_k,x))\nabla_xY_s^\theta(t_k,x)dW_s^j,\end{aligned} \tag{3.4}$$

$$\begin{aligned}\partial_{\theta_i}Y_t^\theta(t_k,x)&=\int_{t_k}^t\Big(\partial_{\theta_i}b(\theta,s,Y_s^\theta(t_k,x))+\nabla_xb(\theta,s,Y_s^\theta(t_k,x))\partial_{\theta_i}Y_s^\theta(t_k,x)\Big)ds \\ &\quad+\sum_{j=1}^d\int_{t_k}^t\nabla_x\sigma_j(s,Y_s^\theta(t_k,x))\partial_{\theta_i}Y_s^\theta(t_k,x)dW_s^j,\end{aligned} \tag{3.5}$$

for $i\in\{1,\ldots,m\}$, where $\sigma_1,...,\sigma_d:\mathbb{R}^d\to\mathbb{R}^d$ denote the columns of the matrix σ.

Moreover, the random variables $Y_t^\theta(t_k,x)$, $\nabla_xY_t^\theta(t_k,x)$, $(\nabla_xY_t^\theta(t_k,x))^{-1}$ and $\partial_{\theta_i}Y_t^\theta(t_k,x)$ belong to $\mathbb{D}^{1,2}$ for any $t\in[t_k,t_{k+1}]$ (see Nualart [19, Section 2.2]). On the other hand, the

Malliavin derivative $D_sY_t^\theta(t_k,x)$ satisfies the following linear equation

$$D_sY_t^\theta(t_k,x)=\sigma(s,Y_s^\theta(t_k,x))+\int_s^t \nabla_x b(\theta,u,Y_u^\theta(t_k,x))D_sY_u^\theta(t_k,x)du$$
$$+\sum_{j=1}^d\int_s^t \nabla_x\sigma_j(u,Y_u^\theta(t_k,x))D_sY_u^\theta(t_k,x)dW_u^j,$$

for $s\le t$ a.e., and $D_sY_t^\theta(t_k,x)=0$ for $s>t$ a.e. By [19, (2.59)], we have that

$$D_sY_t^\theta(t_k,x)=\nabla_xY_t^\theta(t_k,x)(\nabla_xY_s^\theta(t_k,x))^{-1}\sigma(s,Y_s^\theta(t_k,x))\mathbf{1}_{[t_k,t]}(s).$$

Now, for all $k\in\{0,...,n-1\}$ and $x\in\mathbb{R}^d$, we denote by $\widetilde{\mathrm{P}}_{t_k,x}^\theta$ the probability law of Y^θ starting at x at time t_k, i.e., $\widetilde{\mathrm{P}}_{t_k,x}^\theta(A)=\widetilde{\mathrm{E}}[\mathbf{1}_A|Y_{t_k}^\theta=x]$ for all $A\in\widetilde{\mathcal{F}}$, and denote by $\widetilde{\mathrm{E}}_{t_k,x}^\theta$ the expectation w.r.t. $\widetilde{\mathrm{P}}_{t_k,x}^\theta$. That is, for all $\widetilde{\mathcal{F}}$-measurable random variables V, we have that $\widetilde{\mathrm{E}}_{t_k,x}^\theta[V]=\widetilde{\mathrm{E}}[V|Y_{t_k}^\theta=x]$. Hence, $\widetilde{\mathrm{E}}_{t_k,x}^\theta$ is the expectation under the probability law of Y^θ starting at x at time t_k.

Similarly, we denote by $\widehat{\mathrm{P}}_{t_k,x}^\theta$ the probability law of X^θ starting at x at time t_k, i.e., $\widehat{\mathrm{P}}_{t_k,x}^\theta(A)=\widehat{\mathrm{E}}[\mathbf{1}_A|X_{t_k}^\theta=x]$ for all $A\in\widehat{\mathcal{F}}$, and denote by $\widehat{\mathrm{E}}_{t_k,x}^\theta$ the expectation w.r.t. $\widehat{\mathrm{P}}_{t_k,x}^\theta$. That is, for all $\widehat{\mathcal{F}}$-measurable random variables V, we have that $\widehat{\mathrm{E}}_{t_k,x}^\theta[V]=\widehat{\mathrm{E}}[V|X_{t_k}^\theta=x]$. Let $\mathrm{P}_{t_k,x}^\theta:=\widehat{\mathrm{P}}_{t_k,x}^\theta\otimes\widetilde{\mathrm{P}}_{t_k,x}^\theta$ be the product measure, and $\mathrm{E}_{t_k,x}^\theta=\widehat{\mathrm{E}}_{t_k,x}^\theta\otimes\widetilde{\mathrm{E}}_{t_k,x}^\theta$ denotes the expectation w.r.t. $\mathrm{P}_{t_k,x}^\theta$.

As a consequence of [6, Proposition 4.1], we have the following expression for the logarithm derivative of the transition density w.r.t. θ in terms of a conditional expectation involving Skorohod integral.

Lemma 3.1. *Under conditions* **(A1)**, **(A2)** *and* **(A3)**(a)-(b), *for all $i\in\{1,\dots,m\}$, $k\in\{0,...,n-1\}$, $\theta\in\Theta$, and $x,y\in\mathbb{R}^d$,*

$$\frac{\partial_{\theta_i}p^\theta}{p^\theta}(t_k,t_{k+1},x,y)=\frac{1}{\Delta_n}\widetilde{\mathrm{E}}_{t_k,x}^\theta\left[\delta\left(U^\theta(t_k,x)\partial_{\theta_i}Y_{t_{k+1}}^\theta(t_k,x)\right)|Y_{t_{k+1}}^\theta=y\right],$$

where $U_t^\theta(t_k,x)=(D_tY_{t_{k+1}}^\theta(t_k,x))^{-1}$, $t\in[t_k,t_{k+1}]$.

Now, we have the following decomposition of the Skorohod integral appearing in the conditional expectation of Lemma 3.1.

Lemma 3.2. *Under conditions* **(A1)**, **(A2)** *and* **(A3)**(a)-(b), *for all $i\in\{1,\dots,m\}$, $k\in\{0,...,n-1\}$, $\theta\in\Theta$, and $x\in\mathbb{R}^d$,*

$$\delta\left(U^\theta(t_k,x)\partial_{\theta_i}Y_{t_{k+1}}^\theta(t_k,x)\right)=\Delta_n(\partial_{\theta_i}b(\theta,t_k,x))^*(\sigma\sigma^*)^{-1}(t_k,x)\left(Y_{t_{k+1}}^\theta-Y_{t_k}^\theta-b(\theta,t_k,Y_{t_k}^\theta)\Delta_n\right)$$
$$-R_1^{\theta,k}+R_2^{\theta,k}+R_3^{\theta,k}-R_4^{\theta,k}-R_5^{\theta,k},$$

where

$$R_1^{\theta,k}=\int_{t_k}^{t_{k+1}}\int_s^{t_{k+1}}\mathrm{tr}\left(D_s\big(((\nabla_xY_u^\theta(t_k,x))^{-1}\partial_{\theta_i}b(\theta,u,Y_u^\theta(t_k,x)))^*\big)\sigma^{-1}(s,Y_s^\theta(t_k,x))\nabla_xY_s^\theta(t_k,x)\right)duds,$$

$$R_2^{\theta,k}=\int_{t_k}^{t_{k+1}}((\nabla_xY_s^\theta(t_k,x))^{-1}\partial_{\theta_i}b(\theta,s,Y_s^\theta(t_k,x)))^*ds$$

$$\cdot \int_{t_k}^{t_{k+1}} \Big((\nabla_x Y_s^\theta(t_k,x))^* (\sigma^{-1}(s, Y_s^\theta(t_k,x)))^* - (\nabla_x Y_{t_k}^\theta(t_k,x))^* (\sigma^{-1}(t_k, Y_{t_k}^\theta(t_k,x)))^* \Big) dW_s,$$

$$R_3^{\theta,k} = \int_{t_k}^{t_{k+1}} \Big(((\nabla_x Y_s^\theta(t_k,x))^{-1} \partial_{\theta_i} b(\theta, s, Y_s^\theta(t_k,x)))^* - ((\nabla_x Y_{t_k}^\theta(t_k,x))^{-1} \partial_{\theta_i} b(\theta, t_k, Y_{t_k}^\theta(t_k,x)))^* \Big) ds$$

$$\cdot \int_{t_k}^{t_{k+1}} (\nabla_x Y_{t_k}^\theta(t_k,x))^* (\sigma^{-1}(t_k, Y_{t_k}^\theta(t_k,x)))^* dW_s,$$

$$R_4^{\theta,k} = \Delta_n (\partial_{\theta_i} b(\theta, t_k, Y_{t_k}^\theta))^* (\sigma\sigma^*)^{-1}(t_k, Y_{t_k}^\theta) \int_{t_k}^{t_{k+1}} \Big(b(\theta, s, Y_s^\theta) - b(\theta, t_k, Y_{t_k}^\theta) \Big) ds,$$

$$R_5^{\theta,k} = \Delta_n (\partial_{\theta_i} b(\theta, t_k, Y_{t_k}^\theta))^* (\sigma\sigma^*)^{-1}(t_k, Y_{t_k}^\theta) \int_{t_k}^{t_{k+1}} \Big(\sigma(s, Y_s^\theta) - \sigma(t_k, Y_{t_k}^\theta) \Big) dW_s.$$

As a consequence of Lemmas 3.1 and 3.2, we have the following explicit expression for the logarithm derivative of the transition density w.r.t. the parameter.

Lemma 3.3. *Under conditions* **(A1)**, **(A2)** *and* **(A3)**(a)-(b), *for all* $i \in \{1, \dots, m\}$, $k \in \{0, ..., n-1\}$, $\theta \in \Theta$, *and* $x, y \in \mathbb{R}^d$,

$$\frac{\partial_{\theta_i} p^\theta}{p^\theta}(t_k, t_{k+1}, x, y) = (\partial_{\theta_i} b(\theta, t_k, x))^* (\sigma\sigma^*)^{-1}(t_k, x)\, (y - x - b(\theta, t_k, x)\Delta_n) + \frac{1}{\Delta_n} \widetilde{\mathrm{E}}_{t_k,x}^\theta \left[-R_1^{\theta,k} + R_2^{\theta,k} + R_3^{\theta,k} - R_4^{\theta,k} - R_5^{\theta,k} \big| Y_{t_{k+1}}^\theta = y \right].$$

We will use the following estimates for the solution to (3.2).

Lemma 3.4. *Assume conditions* **(A1)**-**(A2)**.

(i) *For any* $p \geq 1$ *and* $\theta \in \Theta$, *there exists a constant* $C_p > 0$ *such that for all* $k \in \{0, ..., n-1\}$ *and* $t \in [t_k, t_{k+1}]$,

$$\widetilde{\mathrm{E}}_{t_k,x}^\theta \left[\left| Y_t^\theta(t_k,x) - Y_{t_k}^\theta(t_k,x) \right|^p \right] \leq C_p \left| t - t_k \right|^{\frac{p}{2}} (1 + |x|^p).$$

(ii) *For any function* g *defined on* $\Theta \times \mathbb{R}^d$ *with polynomial growth in* x *uniformly in* $\theta \in \Theta$, *there exist constants* $C, q > 0$ *such that for all* $k \in \{0, ..., n-1\}$ *and* $t \in [t_k, t_{k+1}]$,

$$\widetilde{\mathrm{E}}_{t_k,x}^\theta \left[\left| g(\theta, Y_t^\theta(t_k,x)) \right| \right] \leq C (1 + |x|^q).$$

Moreover, all these statements remain valid for X^θ.

Assuming conditions **(A1)**, **(A2)**, and **(A3)**(a)-(b), and using Gronwall's inequality, one can easily check that for any $\theta \in \Theta$ and $p \geq 2$, there exists a constant $C_p > 0$ such that for all $k \in \{0, ..., n-1\}$ and $t \in [t_k, t_{k+1}]$,

$$\widetilde{\mathrm{E}}_{t_k,x}^\theta \left[\left| \nabla_x Y_t^\theta(t_k,x) \right|^p + \left| (\nabla_x Y_t^\theta(t_k,x))^{-1} \right|^p \right] + \sup_{s \in [t_k, t_{k+1}]} \widetilde{\mathrm{E}}_{t_k,x}^\theta \left[\left| D_s Y_t^\theta(t_k,x) \right|^p \right] + \sup_{s \in [t_k, t_{k+1}]} \widetilde{\mathrm{E}}_{t_k,x}^\theta \left[\left| D_s \left(\nabla_x Y_t^\theta(t_k,x) \right) \right|^p \right] \leq C_p, \tag{3.6}$$

$$\widetilde{\mathrm{E}}_{t_k,x}^\theta \left[\left| \partial_{\theta_i} Y_t^\theta(t_k,x) \right|^p \right] \leq C_p (1 + |x|^p), \tag{3.7}$$

where the constant C_p is uniform in θ. As a consequence, we have the following estimates, which follow easily from (6.2), Lemma 3.4 and properties of the moments of the Brownian motion.

Lemma 3.5. *Under conditions* **(A1)**, **(A2)**, *and* **(A3)**(a)-(b)*, for any $\theta \in \Theta$ and $p \geq 2$, there exists a constant $C_p > 0$ such that for all $k \in \{0, ..., n-1\}$,*

$$\widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[-R_1^{\theta,k} + R_2^{\theta,k} + R_3^{\theta,k}\right] = 0, \tag{3.8}$$

$$\widetilde{\mathrm{E}}^{\theta}_{t_k,x}\left[\left|-R_1^{\theta,k} + R_2^{\theta,k} + R_3^{\theta,k}\right|^p\right] \leq C_p\Delta_n^{2p}\left(1 + |x|^p\right). \tag{3.9}$$

We next recall Girsanov's theorem on each interval $[t_k, t_{k+1}]$. For all $\theta, \theta^1 \in \Theta$, $x \in \mathbb{R}^d$ and $k \in \{0, ..., n-1\}$, the probability measures $\widehat{\mathrm{P}}^{\theta}_{t_k,x}$ and $\widehat{\mathrm{P}}^{\theta^1}_{t_k,x}$ are absolutely continuous w.r.t. each other and its Radon-Nikodym derivative is given by

$$\begin{aligned}\frac{d\widehat{\mathrm{P}}^{\theta}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta^1}_{t_k,x}}\left((X_t^{\theta^1})_{t\in[t_k,t_{k+1}]}\right) &= \exp\left\{\int_{t_k}^{t_{k+1}} \sigma^{-1}(t, X_t^{\theta^1})\left(b(\theta, t, X_t^{\theta^1}) - b(\theta^1, t, X_t^{\theta^1})\right)\cdot dB_t\right.\\ &\left.\quad - \frac{1}{2}\int_{t_k}^{t_{k+1}}\left|\sigma^{-1}(t, X_t^{\theta^1})\left(b(\theta, t, X_t^{\theta^1}) - b(\theta^1, t, X_t^{\theta^1})\right)\right|^2 dt\right\}.\end{aligned} \tag{3.10}$$

See [13, Theorem 1.12]. By Girsanov's theorem, the process $B^{\widehat{\mathrm{P}}^{\theta}_{t_k,x}} = (B_t^{\widehat{\mathrm{P}}^{\theta}_{t_k,x}}, t \in [t_k, t_{k+1}])$ is a Brownian motion under $\widehat{\mathrm{P}}^{\theta}_{t_k,x}$, where for any $t \in [t_k, t_{k+1}]$,

$$B_t^{\widehat{\mathrm{P}}^{\theta}_{t_k,x}} := B_t - \int_{t_k}^t \sigma^{-1}(s, X_s^{\theta^1})\left(b(\theta, s, X_s^{\theta^1}) - b(\theta^1, s, X_s^{\theta^1})\right)ds.$$

Next, we give two following technical lemmas which will be useful in the sequel.

Lemma 3.6. *Assume conditions* **(A1)**, **(A2)** *and* **(A3)**(a)*. Let $\theta_0, \theta \in \Theta$. Then for any $k \in \{0, ..., n-1\}$ and $\widetilde{\mathcal{F}}_{t_{k+1}}$-measurable random variable V,*

$$\widehat{\mathrm{E}}^{\theta^0}\left[\widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[V\big|Y^{\theta}_{t_{k+1}} = X^{\theta^0}_{t_{k+1}}\right]\big|\widehat{\mathcal{F}}_{t_k}\right] = \widetilde{\mathrm{E}}^{\theta}_{t_k,X^{\theta^0}_{t_k}}\left[V\right].$$

Now, to simplify the notation, for $j \in \{1, \ldots, m\}$ we set

$$\theta_j(0+) := (\theta_1^0, \ldots, \theta_{j-1}^0, \theta_j, \theta_{j+1}^0 + \varphi_{n\Delta_n}^{j+1}(\theta^0)u_{j+1}, \ldots, \theta_m^0 + \varphi_{n\Delta_n}^m(\theta^0)u_m).$$

Lemma 3.7. *Assume conditions* **(A1)** *and* **(A2)**. *Let $p, q > 1$ satisfying that $\frac{1}{p} + \frac{1}{q} = 1$. Then for any $k \in \{0, ..., n-1\}$ and $x \in \mathbb{R}^d$, there exists a constant $C > 0$ which does not depend on x such that for any $\widehat{\mathcal{F}}_{t_{k+1}}$-measurable random variable V,*

$$\begin{aligned}&\left|\widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,x}\left[V\left(\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\left((X_t^{\theta_i^{0+}(\ell)})_{t\in[t_k,t_{k+1}]}\right) - 1\right)\right]\right|\\ &\leq C\sqrt{\Delta_n}\left(1 + |x|\right)\left(\left|\int_{\theta_i^0+\ell\varphi^i_{n\Delta_n}(\theta^0)u_i}^{\theta_i^0}\left(\widehat{\mathrm{E}}^{\theta_i(0+)}_{t_k,x}\left[|V|^q\right]\right)^{\frac{1}{q}}d\theta_i\right|\right.\\ &\quad + \left|\int_{\theta_{i+1}^0+\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}}^{\theta_{i+1}^0}\left(\widehat{\mathrm{E}}^{\theta_{i+1}(0+)}_{t_k,x}\left[|V|^q\right]\right)^{\frac{1}{q}}d\theta_{i+1}\right|\end{aligned}$$

$$+\cdots+\left|\int_{\theta_m^0+\varphi_{n\Delta_n}^m(\theta^0)u_m}^{\theta_m^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_m(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}d\theta_m\right|\Bigg).$$

We finally recall a convergence in probability result. For each $n\in\mathbb{N}$, let $(\zeta_{k,n})_{k\geq 1}$ be a sequence of random variables defined on the filtered probability space $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},\mathrm{P})$, and assume that they are $\mathcal{F}_{t_{k+1}}$-measurable for all k.

Lemma 3.8. [11, Lemma 3.4] a) *Assume that as $n\to\infty$,*

$$\text{(i)}\ \sum_{k=0}^{n-1}\mathrm{E}\left[\zeta_{k,n}|\mathcal{F}_{t_k}\right]\xrightarrow{\mathrm{P}}0,\qquad and\qquad \text{(ii)}\ \sum_{k=0}^{n-1}\mathrm{E}\left[\zeta_{k,n}^2|\mathcal{F}_{t_k}\right]\xrightarrow{\mathrm{P}}0.$$

Then as $n\to\infty$, $\sum_{k=0}^{n-1}\zeta_{k,n}\xrightarrow{\mathrm{P}}0$.

b) *Assume that $\sum_{k=0}^{n-1}\mathrm{E}\left[|\zeta_{k,n}||\mathcal{F}_{t_k}\right]\xrightarrow{\mathrm{P}}0$ as $n\to\infty$. Then as $n\to\infty$, $\sum_{k=0}^{n-1}\zeta_{k,n}\xrightarrow{\mathrm{P}}0$.*

4. Proof of Theorem 2.1

In this section, the proof of Theorem 2.1 will be divided into three steps. We begin deriving an appropriate stochastic expansion of the log-likelihood ratio using Lemma 3.3. The second step deals with the main contributions by applying the multivariate central limit theorem for continuous local martingales in order to show the LAMN property. Finally, the last step is devoted to treat the negligible contributions of the expansion.

4.1. **Expansion of the log-likelihood ratio.**

Lemma 4.1. *Assume conditions* **(A1)**, **(A2)** *and* **(A3)**(a)-(b). *Then*

$$\log\frac{d\mathrm{P}_n^{\theta^0+\varphi_{n\Delta_n}(\theta^0)u}}{d\mathrm{P}_n^{\theta^0}}\left(X^{n,\theta^0}\right)=\sum_{k=0}^{n-1}\sum_{i=1}^m\xi_{i,k,n}+\sum_{k=0}^{n-1}\sum_{i=1}^m\frac{\varphi_{n\Delta_n}^i(\theta^0)u_i}{\Delta_n}\int_0^1\Bigg\{Z_{i,k,n}^{4,\ell}+Z_{i,k,n}^{5,\ell}$$
$$+\widetilde{\mathrm{E}}_{t_k,X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}\left[R^{\theta_i^{0+}(\ell),k}-R_4^{\theta_i^{0+}(\ell),k}-R_5^{\theta_i^{0+}(\ell),k}\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X_{t_{k+1}}^{\theta^0}\right]\Bigg\}d\ell,$$

where

$$\xi_{i,k,n}=\varphi_{n\Delta_n}^i(\theta^0)u_i\int_0^1(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(t_k,X_{t_k}^{\theta^0})$$
$$\cdot\left(\sigma(t_k,X_{t_k}^{\theta^0})\left(B_{t_{k+1}}-B_{t_k}\right)+\left(b(\theta^0,t_k,X_{t_k}^{\theta^0})-b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0})\right)\Delta_n\right)d\ell,$$

$$R^{\theta_i^{0+}(\ell),k}=-R_1^{\theta_i^{0+}(\ell),k}+R_2^{\theta_i^{0+}(\ell),k}+R_3^{\theta_i^{0+}(\ell),k},$$

$$Z_{i,k,n}^{4,\ell}=\Delta_n(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(t_k,X_{t_k}^{\theta^0})\int_{t_k}^{t_{k+1}}\left(b(\theta^0,s,X_s^{\theta^0})-b(\theta^0,t_k,X_{t_k}^{\theta^0})\right)ds,$$

$$Z_{i,k,n}^{5,\ell}=\Delta_n(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(t_k,X_{t_k}^{\theta^0})\int_{t_k}^{t_{k+1}}\left(\sigma(s,X_s^{\theta^0})-\sigma(t_k,X_{t_k}^{\theta^0})\right)dB_s.$$

Proof. Using the decomposition (3.1) and Lemma 3.3, we obtain that

$$\log\frac{d\mathrm{P}_n^{\theta^0+\varphi_{n\Delta_n}(\theta^0)u}}{d\mathrm{P}_n^{\theta^0}}\left(X^{n,\theta^0}\right)=\sum_{k=0}^{n-1}\sum_{i=1}^m\varphi_{n\Delta_n}^i(\theta^0)u_i\int_0^1\frac{\partial_{\theta_i}p^{\theta_i^{0+}(\ell)}}{p^{\theta_i^{0+}(\ell)}}\left(t_k,t_{k+1},X_{t_k}^{\theta^0},X_{t_{k+1}}^{\theta^0}\right)d\ell$$

$$= \sum_{k=0}^{n-1}\sum_{i=1}^{m} \varphi^i_{n\Delta_n}(\theta^0) u_i \int_0^1 \Big((\partial_{\theta_i} b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0}))^* (\sigma\sigma^*)^{-1}(t_k, X_{t_k}^{\theta^0}) \Big(X_{t_{k+1}}^{\theta^0} - X_{t_k}^{\theta^0} \tag{4.1}$$
$$- b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0})\Delta_n \Big) + \frac{1}{\Delta_n} \widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k, X_{t_k}^{\theta^0}} \Big[R^{\theta_i^{0+}(\ell),k} - R_4^{\theta_i^{0+}(\ell),k} - R_5^{\theta_i^{0+}(\ell),k} \Big| Y_{t_{k+1}}^{\theta_i^{0+}(\ell)} = X_{t_{k+1}}^{\theta^0} \Big] \Big) d\ell.$$

Next, using equation (1.1), we get that

$$X_{t_{k+1}}^{\theta^0} - X_{t_k}^{\theta^0} = \sigma(t_k, X_{t_k}^{\theta^0}) \left(B_{t_{k+1}} - B_{t_k} \right) + b(\theta^0, t_k, X_{t_k}^{\theta^0})\Delta_n$$
$$+ \int_{t_k}^{t_{k+1}} \Big(b(\theta^0, s, X_s^{\theta^0}) - b(\theta^0, t_k, X_{t_k}^{\theta^0}) \Big) ds + \int_{t_k}^{t_{k+1}} \Big(\sigma(s, X_s^{\theta^0}) - \sigma(t_k, X_{t_k}^{\theta^0}) \Big) dB_s.$$

This, together with (4.1), gives the desired result. □

In the next two subsections, we will show that $\xi_{i,k,n}$ is the only term that contributes to the limit and all the others terms are negligible.

4.2. Main contributions: LAMN property.

Lemma 4.2. *Assume conditions* **(A1)**- **(A6)**. *Then, as* $n \to \infty$,

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m} \xi_{i,k,n} \overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})}{\longrightarrow} u^*\Gamma(\theta^0)^{1/2}\mathcal{N}(0, I_m) - \frac{1}{2} u^* \Gamma(\theta^0) u,$$

where $\mathcal{N}(0, I_m)$ *is independent of* $\Gamma(\theta^0)$.

Proof. Observe that $\xi_{i,k,n} = \xi_{1,i,k,n} + \xi_{2,i,k,n}$, where

$$\xi_{1,i,k,n} = \varphi^i_{n\Delta_n}(\theta^0) u_i \int_0^1 (\partial_{\theta_i} b(\theta^0, t_k, X_{t_k}^{\theta^0}))^* (\sigma\sigma^*)^{-1}(t_k, X_{t_k}^{\theta^0})$$
$$\cdot \Big(\sigma(t_k, X_{t_k}^{\theta^0}) \left(B_{t_{k+1}} - B_{t_k} \right) + \Big(b(\theta^0, t_k, X_{t_k}^{\theta^0}) - b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0}) \Big) \Delta_n \Big) d\ell,$$

$$\xi_{2,i,k,n} = \varphi^i_{n\Delta_n}(\theta^0) u_i \int_0^1 (\partial_{\theta_i} b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0}) - \partial_{\theta_i} b(\theta^0, t_k, X_{t_k}^{\theta^0}))^* (\sigma\sigma^*)^{-1}(t_k, X_{t_k}^{\theta^0})$$
$$\cdot \Big(\sigma(t_k, X_{t_k}^{\theta^0}) \left(B_{t_{k+1}} - B_{t_k} \right) + \Big(b(\theta^0, t_k, X_{t_k}^{\theta^0}) - b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0}) \Big) \Delta_n \Big) d\ell.$$

We write $\xi_{1,i,k,n} = \xi_{1,1,i,k,n} + \xi_{1,2,i,k,n}$, where

$$\xi_{1,1,i,k,n} = \varphi^i_{n\Delta_n}(\theta^0) u_i (\partial_{\theta_i} b(\theta^0, t_k, X_{t_k}^{\theta^0}))^* (\sigma\sigma^*)^{-1}(t_k, X_{t_k}^{\theta^0}) \sigma(t_k, X_{t_k}^{\theta^0}) \left(B_{t_{k+1}} - B_{t_k} \right),$$

$$\xi_{1,2,i,k,n} = \varphi^i_{n\Delta_n}(\theta^0) u_i \int_0^1 (\partial_{\theta_i} b(\theta^0, t_k, X_{t_k}^{\theta^0}))^* (\sigma\sigma^*)^{-1}(t_k, X_{t_k}^{\theta^0})$$
$$\cdot \Big(b(\theta^0, t_k, X_{t_k}^{\theta^0}) - b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta^0}) \Big) \Delta_n d\ell.$$

First, notice that

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m} \xi_{1,1,i,k,n} = \sum_{k=0}^{n-1}\sum_{i=1}^{m} \varphi^i_{n\Delta_n}(\theta^0) u_i (\partial_{\theta_i} b(\theta^0, t_k, X_{t_k}^{\theta^0}))^* (\sigma\sigma^*)^{-1}(t_k, X_{t_k}^{\theta^0}) \sigma(t_k, X_{t_k}^{\theta^0}) \left(B_{t_{k+1}} - B_{t_k} \right)$$
$$= \sum_{k=0}^{n-1} u^* \varphi_{n\Delta_n}(\theta^0) \int_{t_k}^{t_{k+1}} \sigma^{-1}(t_k, X_{t_k}^{\theta^0}) \nabla_\theta b(\theta^0, t_k, X_{t_k}^{\theta^0}) dB_t$$

$$= u^*\varphi_{n\Delta_n}(\theta^0)\int_0^{n\Delta_n}\sigma^{-1}(t,X_t^{\theta^0})\nabla_\theta b(\theta^0,t,X_t^{\theta^0})dB_t - \sum_{k=0}^{n-1}H_{1,k,n},$$

where

$$H_{1,k,n} = u^*\varphi_{n\Delta_n}(\theta^0)\int_{t_k}^{t_{k+1}}\Big(\sigma^{-1}(t,X_t^{\theta^0})\nabla_\theta b(\theta^0,t,X_t^{\theta^0}) - \sigma^{-1}(t_k,X_{t_k}^{\theta^0})\nabla_\theta b(\theta^0,t_k,X_{t_k}^{\theta^0})\Big)dB_t.$$

Next, we treat $\xi_{1,2,i,k,n}$. For this, using the mean value theorem,

$$\begin{aligned}
&b(\theta^0,t_k,X_{t_k}^{\theta^0}) - b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0}) = -\Big(b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0}) - b(\theta_{i+1}^{0+},t_k,X_{t_k}^{\theta^0})\\
&\quad + b(\theta_{i+1}^{0+},t_k,X_{t_k}^{\theta^0}) - b(\theta_{i+2}^{0+},t_k,X_{t_k}^{\theta^0}) + \cdots + b(\theta_{m-1}^{0+},t_k,X_{t_k}^{\theta^0}) - b(\theta_m^{0+},t_k,X_{t_k}^{\theta^0})\\
&\quad + b(\theta_m^{0+},t_k,X_{t_k}^{\theta^0}) - b(\theta^0,t_k,X_{t_k}^{\theta^0})\Big)\\
&= -\Big(\ell\varphi_{n\Delta_n}^i(\theta^0)u_i\int_0^1\partial_{\theta_i}b(\theta_i^{0+}(\alpha\ell),t_k,X_{t_k}^{\theta^0})d\alpha + \varphi_{n\Delta_n}^{i+1}(\theta^0)u_{i+1}\int_0^1\partial_{\theta_{i+1}}b(\theta_{i+1}^{0+}(\alpha),t_k,X_{t_k}^{\theta^0})d\alpha\\
&\quad + \cdots + \varphi_{n\Delta_n}^m(\theta^0)u_m\int_0^1\partial_{\theta_m}b(\theta_m^{0+}(\alpha),t_k,X_{t_k}^{\theta^0})d\alpha\Big)\\
&= -\Big(\ell\varphi_{n\Delta_n}^i(\theta^0)u_i\partial_{\theta_i}b(\theta^0,t_k,X_{t_k}^{\theta^0}) + \varphi_{n\Delta_n}^{i+1}(\theta^0)u_{i+1}\partial_{\theta_{i+1}}b(\theta^0,t_k,X_{t_k}^{\theta^0})\\
&\quad + \cdots + \varphi_{n\Delta_n}^m(\theta^0)u_m\partial_{\theta_m}b(\theta^0,t_k,X_{t_k}^{\theta^0})d\alpha\Big)\\
&\quad - \Big(\ell\varphi_{n\Delta_n}^i(\theta^0)u_i\int_0^1\Big(\partial_{\theta_i}b(\theta_i^{0+}(\alpha\ell),t_k,X_{t_k}^{\theta^0}) - \partial_{\theta_i}b(\theta^0,t_k,X_{t_k}^{\theta^0})\Big)d\alpha\\
&\quad + \varphi_{n\Delta_n}^{i+1}(\theta^0)u_{i+1}\int_0^1\Big(\partial_{\theta_{i+1}}b(\theta_{i+1}^{0+}(\alpha),t_k,X_{t_k}^{\theta^0}) - \partial_{\theta_{i+1}}b(\theta^0,t_k,X_{t_k}^{\theta^0})\Big)d\alpha\\
&\quad + \cdots + \varphi_{n\Delta_n}^m(\theta^0)u_m\int_0^1\Big(\partial_{\theta_m}b(\theta_m^{0+}(\alpha),t_k,X_{t_k}^{\theta^0}) - \partial_{\theta_m}b(\theta^0,t_k,X_{t_k}^{\theta^0})\Big)d\alpha\Big),
\end{aligned}$$

where, to simplify the exposition, we have set for $j \in \{i+1,\ldots,m\}$,

$$\begin{aligned}
\theta_i^{0+}(\alpha\ell) &:= (\theta_1^0,\ldots,\theta_{i-1}^0,\theta_i^0+\alpha\ell\varphi_{n\Delta_n}^i(\theta^0)u_i,\theta_{i+1}^0+\varphi_{n\Delta_n}^{i+1}(\theta^0)u_{i+1},\ldots,\theta_m^0+\varphi_{n\Delta_n}^m(\theta^0)u_m),\\
\theta_j^{0+}(\alpha) &:= (\theta_1^0,\ldots,\theta_{j-1}^0,\theta_j^0+\alpha\varphi_{n\Delta_n}^j(\theta^0)u_j,\theta_{j+1}^0+\varphi_{n\Delta_n}^{j+1}(\theta^0)u_{j+1},\ldots,\theta_m^0+\varphi_{n\Delta_n}^m(\theta^0)u_m).
\end{aligned}$$

This implies that

$$\begin{aligned}
&\sum_{k=0}^{n-1}\sum_{i=1}^m\xi_{1,2,i,k,n} = \sum_{k=0}^{n-1}\sum_{i=1}^m\varphi_{n\Delta_n}^i(\theta^0)u_i\int_0^1(\partial_{\theta_i}b(\theta^0,t_k,X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(t_k,X_{t_k}^{\theta^0})\\
&\quad\cdot\Big(b(\theta^0,t_k,X_{t_k}^{\theta^0}) - b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0})\Big)\Delta_n d\ell\\
&= -\sum_{k=0}^{n-1}\sum_{i=1}^m\varphi_{n\Delta_n}^i(\theta^0)u_i(\partial_{\theta_i}b(\theta^0,t_k,X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(t_k,X_{t_k}^{\theta^0})\\
&\quad\cdot\Big(\frac{1}{2}\varphi_{n\Delta_n}^i(\theta^0)u_i\partial_{\theta_i}b(\theta^0,t_k,X_{t_k}^{\theta^0}) + \varphi_{n\Delta_n}^{i+1}(\theta^0)u_{i+1}\partial_{\theta_{i+1}}b(\theta^0,t_k,X_{t_k}^{\theta^0})
\end{aligned}$$

$$+\cdots+\varphi^m_{n\Delta_n}(\theta^0)u_m\partial_{\theta_m}b(\theta^0,t_k,X^{\theta^0}_{t_k})\Big)\Delta_n-\sum_{k=0}^{n-1}\sum_{i=1}^{m}\left(K_{i,k,n}+K_{i+1,k,n}+\cdots+K_{m,k,n}\right)$$

$$=-\frac{1}{2}\sum_{k=0}^{n-1}u^*\varphi_{n\Delta_n}(\theta^0)\int_{t_k}^{t_{k+1}}(\nabla_\theta b(\theta^0,t_k,X^{\theta^0}_{t_k}))^*(\sigma^{-1}(t_k,X^{\theta^0}_{t_k}))^*\sigma^{-1}(t_k,X^{\theta^0}_{t_k})\nabla_\theta b(\theta^0,t_k,X^{\theta^0}_{t_k})dt\varphi_{n\Delta_n}(\theta^0)u$$
$$-\sum_{k=0}^{n-1}\sum_{i=1}^{m}\left(K_{i,k,n}+K_{i+1,k,n}+\cdots+K_{m,k,n}\right)$$
$$=-\frac{1}{2}u^*\varphi_{n\Delta_n}(\theta^0)\int_0^{n\Delta_n}(\nabla_\theta b(\theta^0,t,X^{\theta^0}_t))^*(\sigma^{-1}(t,X^{\theta^0}_t))^*\sigma^{-1}(t,X^{\theta^0}_t)\nabla_\theta b(\theta^0,t,X^{\theta^0}_t)dt\varphi_{n\Delta_n}(\theta^0)u$$
$$+\frac{1}{2}\sum_{k=0}^{n-1}H_{2,k,n}-\sum_{k=0}^{n-1}\sum_{i=1}^{m}\left(K_{i,k,n}+K_{i+1,k,n}+\cdots+K_{m,k,n}\right),$$

where for $j\in\{i+1,\dots,m\}$,

$$K_{i,k,n}=\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1\int_0^1(\partial_{\theta_i}b(\theta^0,t_k,X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k,X^{\theta^0}_{t_k})$$
$$\cdot\,\ell\varphi^i_{n\Delta_n}(\theta^0)u_i\Big(\partial_{\theta_i}b(\theta_i^{0+}(\alpha\ell),t_k,X^{\theta^0}_{t_k})-\partial_{\theta_i}b(\theta^0,t_k,X^{\theta^0}_{t_k})\Big)\Delta_n d\alpha d\ell,$$
$$K_{j,k,n}=\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1(\partial_{\theta_i}b(\theta^0,t_k,X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k,X^{\theta^0}_{t_k})$$
$$\cdot\,\varphi^j_{n\Delta_n}(\theta^0)u_j\Big(\partial_{\theta_j}b(\theta_j^{0+}(\alpha),t_k,X^{\theta^0}_{t_k})-\partial_{\theta_j}b(\theta^0,t_k,X^{\theta^0}_{t_k})\Big)\Delta_n d\alpha,$$
$$H_{2,k,n}=u^*\varphi_{n\Delta_n}(\theta^0)\int_{t_k}^{t_{k+1}}\Big((\nabla_\theta b(\theta^0,t,X^{\theta^0}_t))^*(\sigma^{-1}(t,X^{\theta^0}_t))^*\sigma^{-1}(t,X^{\theta^0}_t)\nabla_\theta b(\theta^0,t,X^{\theta^0}_t)$$
$$-(\nabla_\theta b(\theta^0,t_k,X^{\theta^0}_{t_k}))^*(\sigma^{-1}(t_k,X^{\theta^0}_{t_k}))^*\sigma^{-1}(t_k,X^{\theta^0}_{t_k})\nabla_\theta b(\theta^0,t_k,X^{\theta^0}_{t_k})\Big)dt\;\varphi_{n\Delta_n}(\theta^0)u.$$

Therefore, we have shown that

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m}\xi_{i,k,n}=\sum_{k=0}^{n-1}\sum_{i=1}^{m}\left(\xi_{1,i,k,n}+\xi_{2,i,k,n}\right)=\sum_{k=0}^{n-1}\sum_{i=1}^{m}\left(\xi_{1,1,i,k,n}+\xi_{1,2,i,k,n}+\xi_{2,i,k,n}\right)$$
$$=u^*\varphi_{n\Delta_n}(\theta^0)\int_0^{n\Delta_n}\sigma^{-1}(t,X^{\theta^0}_t)\nabla_\theta b(\theta^0,t,X^{\theta^0}_t)dB_t$$
$$-\frac{1}{2}u^*\varphi_{n\Delta_n}(\theta^0)\int_0^{n\Delta_n}(\nabla_\theta b(\theta^0,t,X^{\theta^0}_t))^*(\sigma^{-1}(t,X^{\theta^0}_t))^*\sigma^{-1}(t,X^{\theta^0}_t)\nabla_\theta b(\theta^0,t,X^{\theta^0}_t)dt\varphi_{n\Delta_n}(\theta^0)u$$
$$-\sum_{k=0}^{n-1}H_{1,k,n}+\frac{1}{2}\sum_{k=0}^{n-1}H_{2,k,n}-\sum_{k=0}^{n-1}\sum_{i=1}^{m}\left(K_{i,k,n}+K_{i+1,k,n}+\cdots+K_{m,k,n}\right)+\sum_{k=0}^{n-1}\sum_{i=1}^{m}\xi_{2,i,k,n}. \tag{4.2}$$

Next, using condition **(A4)**, as $n\to\infty$,

$$\varphi_{n\Delta_n}(\theta^0)\int_0^{n\Delta_n}(\nabla_\theta b(\theta^0,t,X^{\theta^0}_t))^*(\sigma^{-1}(t,X^{\theta^0}_t))^*\sigma^{-1}(t,X^{\theta^0}_t)\nabla_\theta b(\theta^0,t,X^{\theta^0}_t)dt\;\varphi_{n\Delta_n}(\theta^0)\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}\Gamma(\theta^0). \tag{4.3}$$

Then, using the multivariate central limit theorem for continuous local martingales (see [22, Theorem 4.1]), we obtain that as $n \to \infty$,

$$u^* \varphi_{n\Delta_n}(\theta^0) \int_0^{n\Delta_n} \sigma^{-1}(t, X_t^{\theta^0}) \nabla_\theta b(\theta^0, t, X_t^{\theta^0}) dB_t \overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})}{\longrightarrow} u^* \Gamma(\theta^0)^{1/2} \mathcal{N}(0, I_m), \tag{4.4}$$

where $\mathcal{N}(0, I_m)$ is independent of $\Gamma(\theta^0)$. Finally, by Lemma 4.7 below, as $n \to \infty$,

$$\begin{aligned} &-\sum_{k=0}^{n-1} H_{1,k,n} + \frac{1}{2}\sum_{k=0}^{n-1} H_{2,k,n} - \sum_{k=0}^{n-1}\sum_{i=1}^{m} (K_{i,k,n} + K_{i+1,k,n} + \cdots + K_{m,k,n}) \\ &+ \sum_{k=0}^{n-1}\sum_{i=1}^{m} \xi_{2,i,k,n} \overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow} 0. \end{aligned} \tag{4.5}$$

Thus, the desired result follows from (4.2)-(4.5). □

4.3. **Negligible contributions.**

Lemma 4.3. *Under conditions* **(A1)-(A3)** *and* **(A5)-(A6)**, *as* $n \to \infty$,

$$\begin{aligned} &\sum_{k=0}^{n-1}\sum_{i=1}^{m} \frac{\varphi_{n\Delta_n}^i(\theta^0) u_i}{\Delta_n} \int_0^1 \Big\{ Z_{i,k,n}^{4,\ell} + Z_{i,k,n}^{5,\ell} \\ &\quad + \widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)} \left[R^{\theta_i^{0+}(\ell),k} - R_4^{\theta_i^{0+}(\ell),k} - R_5^{\theta_i^{0+}(\ell),k} \Big| Y_{t_{k+1}}^{\theta_i^{0+}(\ell)} = X_{t_{k+1}}^{\theta^0} \right] \Big\} d\ell \overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow} 0. \end{aligned}$$

Proof. The proof is completed by combining three Lemmas 4.4 4.5 and 4.6 below. □

Consequently, from Lemmas 4.1, 4.2 and 4.3, the proof of Theorem 2.1 is now completed.

Lemma 4.4. *Under conditions* **(A1)-(A3)** *and* **(A5)-(A6)**, *as* $n \to \infty$,

$$\sum_{k=0}^{n-1}\sum_{i=1}^{m} \frac{\varphi_{n\Delta_n}^i(\theta^0) u_i}{\Delta_n} \int_0^1 \widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)} \left[R^{\theta_i^{0+}(\ell),k} \Big| Y_{t_{k+1}}^{\theta_i^{0+}(\ell)} = X_{t_{k+1}}^{\theta^0} \right] d\ell \overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow} 0.$$

Proof. It suffices to show that conditions (i) and (ii) of Lemma 3.8 a) hold under the measure $\widehat{\mathrm{P}}^{\theta^0}$ applied to the random variable

$$\zeta_{k,n} = \zeta_{i,k,n} := \frac{\varphi_{n\Delta_n}^i(\theta^0) u_i}{\Delta_n} \int_0^1 \widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)} \left[R^{\theta_i^{0+}(\ell),k} \Big| Y_{t_{k+1}}^{\theta_i^{0+}(\ell)} = X_{t_{k+1}}^{\theta^0} \right] d\ell,$$

for any $i \in \{1, \dots, m\}$. We start showing (i) of Lemma 3.8 a). Applying Lemma 3.6 to $\theta = \theta_i^{0+}(\ell)$ and $V = R^{\theta_i^{0+}(\ell),k}$, and using the fact that, by (3.8), $\widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)}[R^{\theta_i^{0+}(\ell),k}] = 0$, we obtain that

$$\begin{aligned} \sum_{k=0}^{n-1} \widehat{\mathrm{E}}^{\theta^0}\left[\zeta_{i,k,n} | \widehat{\mathcal{F}}_{t_k} \right] &= \sum_{k=0}^{n-1} \frac{\varphi_{n\Delta_n}^i(\theta^0) u_i}{\Delta_n} \int_0^1 \widehat{\mathrm{E}}^{\theta^0}\left[\widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)} \left[R^{\theta_i^{0+}(\ell),k} | Y_{t_{k+1}}^{\theta_i^{0+}(\ell)} = X_{t_{k+1}}^{\theta^0} \right] | \widehat{\mathcal{F}}_{t_k} \right] d\ell \\ &= \sum_{k=0}^{n-1} \frac{\varphi_{n\Delta_n}^i(\theta^0) u_i}{\Delta_n} \int_0^1 \widetilde{\mathrm{E}}_{t_k, X_{t_k}^{\theta^0}}^{\theta_i^{0+}(\ell)} [R^{\theta_i^{0+}(\ell),k}] d\ell = 0. \end{aligned}$$

Thus, the term appearing in condition (i) of Lemma 3.8 a) actually equals zero.

Next, applying Jensen's inequality and Lemma 3.6 to $\theta=\theta_i^{0+}(\ell)$ and $V=(R^{\theta_i^{0+}(\ell),k})^2$, and (3.9), we obtain that

$$\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[\zeta_{i,k,n}^2|\widehat{\mathcal{F}}_{t_k}\right]=\sum_{k=0}^{n-1}\frac{(\varphi^i_{n\Delta_n}(\theta^0))^2u_i^2}{\Delta_n^2}\widehat{\mathrm{E}}^{\theta^0}\left[\left(\int_0^1\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[R^{\theta_i^{0+}(\ell),k}\Big|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]d\ell\right)^2|\widehat{\mathcal{F}}_{t_k}\right]$$

$$\leq\sum_{k=0}^{n-1}\frac{(\varphi^i_{n\Delta_n}(\theta^0))^2u_i^2}{\Delta_n^2}\int_0^1\widehat{\mathrm{E}}^{\theta^0}\left[\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[\left(R^{\theta_i^{0+}(\ell),k}\right)^2|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]|\widehat{\mathcal{F}}_{t_k}\right]d\ell$$

$$=\sum_{k=0}^{n-1}\frac{(\varphi^i_{n\Delta_n}(\theta^0))^2u_i^2}{\Delta_n^2}\int_0^1\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[\left(R^{\theta_i^{0+}(\ell),k}\right)^2\right]d\ell$$

$$\leq C\Delta_n^2(\varphi^i_{n\Delta_n}(\theta^0))^2u_i^2\sum_{k=0}^{n-1}\left(1+|X^{\theta^0}_{t_k}|^2\right)$$

$$\leq Cu_i^2n\Delta_n^2(\varphi^i_{n\Delta_n}(\theta^0))^2+Cu_i^2n\Delta_n^2\max_{k\in\{0,\ldots,n-1\}}|\varphi^i_{t_k}(\theta^0)X^{\theta^0}_{t_k}|^2,$$

which, by conditions **(A5)-(A6)**, converges to zero in $\widehat{\mathrm{P}}^{\theta^0}$-probability as $n\to\infty$. Thus, we have shown that $\sum_{k=0}^{n-1}\zeta_{i,k,n}\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$ for any $i\in\{1,\ldots,m\}$. Thus, the result follows. □

Lemma 4.5. *Under conditions* **(A1)**-**(A3)** *and* **(A5)**-**(A6)***, as* $n\to\infty$*,*

$$\sum_{k=0}^{n-1}\sum_{i=1}^m\frac{\varphi^i_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1\left\{Z^{4,\ell}_{i,k,n}-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[R_4^{\theta_i^{0+}(\ell),k}\Big|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\right\}d\ell\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0.$$

Proof. We rewrite

$$Z^{4,\ell}_{i,k,n}-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[R_4^{\theta_i^{0+}(\ell),k}\Big|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]$$

$$=\Delta_n(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k,X^{\theta^0}_{t_k})\Bigg(\int_{t_k}^{t_{k+1}}\left(b(\theta^0,s,X^{\theta^0}_s)-b(\theta^0,t_k,X^{\theta^0}_{t_k})\right)ds$$

$$-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[\int_{t_k}^{t_{k+1}}\left(b(\theta_i^{0+}(\ell),s,Y^{\theta_i^{0+}(\ell)}_s)-b(\theta_i^{0+}(\ell),t_k,Y^{\theta_i^{0+}(\ell)}_{t_k})\right)ds\Big|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\Bigg)$$

$$=\Delta_n(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k,X^{\theta^0}_{t_k})\left(M_{i,1,k,n}+M_{i,2,k,n}\right),$$

where

$$M_{i,1,k,n}=\int_{t_k}^{t_{k+1}}\left(b(\theta^0,s,X^{\theta^0}_s)-b(\theta^0,t_k,X^{\theta^0}_{t_k})-\left(b(\theta_i^{0+}(\ell),s,X^{\theta^0}_s)-b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k})\right)\right)ds,$$

$$M_{i,2,k,n}=\int_{t_k}^{t_{k+1}}\left(b(\theta_i^{0+}(\ell),s,X^{\theta^0}_s)-b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k})\right)ds$$

$$-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[\int_{t_k}^{t_{k+1}}\left(b(\theta_i^{0+}(\ell),s,Y^{\theta_i^{0+}(\ell)}_s)-b(\theta_i^{0+}(\ell),t_k,Y^{\theta_i^{0+}(\ell)}_{t_k})\right)ds\Big|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right].$$

Thus,

$$\zeta_{k,n}=\zeta_{i,k,n}:=\frac{\varphi^i_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1\left\{Z^{4,\ell}_{i,k,n}-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[R_4^{\theta_i^{0+}(\ell),k}\Big|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\right\}d\ell$$

$$= \frac{\varphi^i_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1 \Delta_n(\partial_{\theta_i} b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k,X^{\theta^0}_{t_k})\left(M_{i,1,k,n}+M_{i,2,k,n}\right)d\ell$$
$$= \zeta_{i,1,k,n}+\zeta_{i,2,k,n},$$

where

$$\zeta_{i,1,k,n} = \varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1 (\partial_{\theta_i} b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k,X^{\theta^0}_{t_k})M_{i,1,k,n}d\ell,$$
$$\zeta_{i,2,k,n} = \varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1 (\partial_{\theta_i} b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k,X^{\theta^0}_{t_k})M_{i,2,k,n}d\ell.$$

Now, using the mean value theorem,

$$\begin{aligned}
&b(\theta^0,s,X_s^{\theta^0})-b(\theta_i^{0+}(\ell),s,X_s^{\theta^0}) = b(\theta^{0+}_{i+1},s,X_s^{\theta^0})-b(\theta_i^{0+}(\ell),s,X_s^{\theta^0})+b(\theta^{0+}_{i+2},s,X_s^{\theta^0})\\
&\quad -b(\theta^{0+}_{i+1},s,X_s^{\theta^0})+\cdots+b(\theta^{0+}_{m},s,X_s^{\theta^0})-b(\theta^{0+}_{m-1},s,X_s^{\theta^0})+b(\theta^0,s,X_s^{\theta^0})-b(\theta^{0+}_{m},s,X_s^{\theta^0})\\
&= -\ell\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1\partial_{\theta_i}b(\theta_i^{0+}(\alpha\ell),s,X_s^{\theta^0})d\alpha-\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}\int_0^1\partial_{\theta_{i+1}}b(\theta^{0+}_{i+1}(\alpha),s,X_s^{\theta^0})d\alpha\\
&\quad -\cdots-\varphi^m_{n\Delta_n}(\theta^0)u_m\int_0^1\partial_{\theta_m}b(\theta_m^{0+}(\alpha),s,X_s^{\theta^0})d\alpha.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&b(\theta^0,s,X_s^{\theta^0})-b(\theta^0,t_k,X_{t_k}^{\theta^0})-\big(b(\theta_i^{0+}(\ell),s,X_s^{\theta^0})-b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0})\big)\\
&= b(\theta^0,s,X_s^{\theta^0})-b(\theta_i^{0+}(\ell),s,X_s^{\theta^0})-\big(b(\theta^0,t_k,X_{t_k}^{\theta^0})-b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0})\big)\\
&= -\ell\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1\Big(\partial_{\theta_i}b(\theta_i^{0+}(\alpha\ell),s,X_s^{\theta^0})-\partial_{\theta_i}b(\theta_i^{0+}(\alpha\ell),t_k,X_{t_k}^{\theta^0})\Big)d\alpha\\
&\quad -\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}\int_0^1\Big(\partial_{\theta_{i+1}}b(\theta^{0+}_{i+1}(\alpha),s,X_s^{\theta^0})-\partial_{\theta_{i+1}}b(\theta^{0+}_{i+1}(\alpha),t_k,X_{t_k}^{\theta^0})\Big)d\alpha\\
&\quad -\cdots-\varphi^m_{n\Delta_n}(\theta^0)u_m\int_0^1\Big(\partial_{\theta_m}b(\theta^{0+}_{m}(\alpha),s,X_s^{\theta^0})-\partial_{\theta_m}b(\theta^{0+}_{m}(\alpha),t_k,X_{t_k}^{\theta^0})\Big)d\alpha.
\end{aligned}$$

Next, using the mean value theorem for vector-valued functions,

$$\begin{aligned}
&\partial_{\theta_j}b(\theta_j^{0+}(\alpha),s,X_s^{\theta^0})-\partial_{\theta_j}b(\theta_j^{0+}(\alpha),t_k,X_{t_k}^{\theta^0})\\
&=\left(\int_0^1 J_{\partial_{\theta_j}b}(t_k+v(s-t_k),X_{t_k}^{\theta^0}+v(X_s^{\theta^0}-X_{t_k}^{\theta^0}))dv\right)\cdot\begin{pmatrix} s-t_k\\ X_s^{\theta^0}-X_{t_k}^{\theta^0}\end{pmatrix},
\end{aligned}$$

for all $j\in\{i,\ldots,m\}$, where the Jacobian matrix is given by

$$\begin{aligned}
&J_{\partial_{\theta_j}b}(t_k+v(s-t_k),X_{t_k}^{\theta^0}+v(X_s^{\theta^0}-X_{t_k}^{\theta^0}))\\
&=\begin{pmatrix}\partial^2_{\theta_j t}b_1 & \partial^2_{\theta_j x_1}b_1 & \ldots & \partial^2_{\theta_j x_d}b_1\\ \vdots & \vdots & \ddots & \vdots\\ \partial^2_{\theta_j t}b_d & \partial^2_{\theta_j x_1}b_d & \ldots & \partial^2_{\theta_j x_d}b_d\end{pmatrix}(\theta_j^{0+}(\alpha),t_k+v(s-t_k),X_{t_k}^{\theta^0}+v(X_s^{\theta^0}-X_{t_k}^{\theta^0})).
\end{aligned}$$

Then, using conditions **(A2)**-**(A3)** and Lemma 3.4 (i), we get that

$$\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[|\zeta_{i,1,k,n}|\,|\widehat{\mathcal{F}}_{t_k}\right]\leq C\Delta_n^{\frac{3}{2}}\varphi^i_{n\Delta_n}(\theta^0)|u_i|\big(\varphi^i_{n\Delta_n}(\theta^0)|u_i|+\varphi^{i+1}_{n\Delta_n}(\theta^0)|u_{i+1}|$$

$$+\cdots+\varphi^{m}_{n\Delta_n}(\theta^0)|u_m|\big)\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|\Big)$$
$$+C\Delta_n^2\varphi^{i}_{n\Delta_n}(\theta^0)|u_i|\big(\varphi^{i}_{n\Delta_n}(\theta^0)|u_i|+\varphi^{i+1}_{n\Delta_n}(\theta^0)|u_{i+1}|+\cdots+\varphi^{m}_{n\Delta_n}(\theta^0)|u_m|\big)\sum_{k=0}^{n-1}\Big(1+|X^{\theta^0}_{t_k}|^2\Big)$$
$$\leq Cn\Delta_n^{\frac{3}{2}}\Big(|u_i|^2(\varphi^{i}_{n\Delta_n}(\theta^0))^2+|u_i||u_{i+1}|\varphi^{i}_{n\Delta_n}(\theta^0)\varphi^{i+1}_{n\Delta_n}(\theta^0)+\cdots+|u_i||u_m|\varphi^{i}_{n\Delta_n}(\theta^0)\varphi^{m}_{n\Delta_n}(\theta^0)\Big)$$
$$+C|u_i|n\Delta_n^{\frac{3}{2}}\Big(\varphi^{i}_{n\Delta_n}(\theta^0)|u_i|+\varphi^{i+1}_{n\Delta_n}(\theta^0)|u_{i+1}|+\cdots+\varphi^{m}_{n\Delta_n}(\theta^0)|u_m|\Big)\max_{k\in\{0,\ldots,n-1\}}|\varphi^{i}_{t_k}(\theta^0)X^{\theta^0}_{t_k}|$$
$$+Cn\Delta_n^2\Big(|u_i|^2(\varphi^{i}_{n\Delta_n}(\theta^0))^2+|u_i||u_{i+1}|\varphi^{i}_{n\Delta_n}(\theta^0)\varphi^{i+1}_{n\Delta_n}(\theta^0)+\cdots+|u_i||u_m|\varphi^{i}_{n\Delta_n}(\theta^0)\varphi^{m}_{n\Delta_n}(\theta^0)\Big)$$
$$+C|u_i|^2n\Delta_n^2\max_{k\in\{0,\ldots,n-1\}}|\varphi^{i}_{t_k}(\theta^0)X^{\theta^0}_{t_k}|^2+C|u_i||u_{i+1}|n\Delta_n^2\max_{k\in\{0,\ldots,n-1\}}\varphi^{i}_{t_k}(\theta^0)\varphi^{i+1}_{t_k}(\theta^0)|X^{\theta^0}_{t_k}|^2$$
$$+\cdots+C|u_i||u_m|n\Delta_n^2\max_{k\in\{0,\ldots,n-1\}}\varphi^{i}_{t_k}(\theta^0)\varphi^{m}_{t_k}(\theta^0)|X^{\theta^0}_{t_k}|^2,$$

which, by conditions **(A5)-(A6)**, converges to zero in $\widehat{\mathrm{P}}^{\theta^0}$-probability as $n\to\infty$. Thus, by Lemma 3.8 b), $\sum_{k=0}^{n-1}\zeta_{i,1,k,n}\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$ for any $i\in\{1,\ldots,m\}$.

Next, using Girsanov's theorem and Lemma 3.6, we get that

$$\widehat{\mathrm{E}}^{\theta^0}\Big[M_{i,2,k,n}|\widehat{\mathcal{F}}_{t_k}\Big]=\widehat{\mathrm{E}}^{\theta^0}\Big[\int_{t_k}^{t_{k+1}}\Big(b(\theta_i^{0+}(\ell),s,X_s^{\theta^0})-b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0})\Big)ds$$
$$-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[\int_{t_k}^{t_{k+1}}\Big(b(\theta_i^{0+}(\ell),s,Y_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t_k,Y_{t_k}^{\theta_i^{0+}(\ell)})\Big)ds\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X^{\theta^0}_{t_{k+1}}\Big]\,|\widehat{\mathcal{F}}_{t_k}\Big]$$
$$=\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\Big[\int_{t_k}^{t_{k+1}}\Big(b(\theta_i^{0+}(\ell),s,X_s^{\theta^0})-b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0})\Big)ds\Big]$$
$$-\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\Big[\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[\int_{t_k}^{t_{k+1}}\Big(b(\theta_i^{0+}(\ell),s,Y_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t_k,Y_{t_k}^{\theta_i^{0+}(\ell)})\Big)ds\Big|Y_{t_{k+1}}^{\theta_i^{0+}(\ell)}=X^{\theta^0}_{t_{k+1}}\Big]\Big]$$
$$=\widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[\int_{t_k}^{t_{k+1}}\Big(b(\theta_i^{0+}(\ell),s,X_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta_i^{0+}(\ell)})\Big)ds\;\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}}\Big]$$
$$-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[\int_{t_k}^{t_{k+1}}\Big(b(\theta_i^{0+}(\ell),s,Y_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t_k,Y_{t_k}^{\theta_i^{0+}(\ell)})\Big)ds\Big]$$
$$=\widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[\int_{t_k}^{t_{k+1}}\big(b(\theta_i^{0+}(\ell),s,X_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta_i^{0+}(\ell)})\big)ds\;\Big(\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}}-1\Big)\Big]$$
$$+\widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[\int_{t_k}^{t_{k+1}}\big(b(\theta_i^{0+}(\ell),s,X_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta_i^{0+}(\ell)})\big)ds\Big]$$
$$-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\Big[\int_{t_k}^{t_{k+1}}\Big(b(\theta_i^{0+}(\ell),s,Y_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t_k,Y_{t_k}^{\theta_i^{0+}(\ell)})\Big)ds\Big]$$

$$= \widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[\int_{t_k}^{t_{k+1}} \left(b(\theta_i^{0+}(\ell),s,X_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta_i^{0+}(\ell)})\right)ds\left(\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}}-1\right)\right],$$

where we have used the fact that $X^{\theta_i^{0+}(\ell)}$ is the independent copy of $Y^{\theta_i^{0+}(\ell)}$. Here, to simplify the exposition, we write $\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}}=\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}}\big((X_t^{\theta_i^{0+}(\ell)})_{t\in[t_k,t_{k+1}]}\big)$.

Then, using Lemma 3.7 with $q=2$, conditions **(A1)-(A2)** and Lemma 3.4 (i), we get that

$$\begin{aligned}
&\left|\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[\zeta_{i,2,k,n}|\widehat{\mathcal{F}}_{t_k}\right]\right|\\
&=\left|\sum_{k=0}^{n-1}\varphi^i_{n\Delta_n}(\theta^0)u_i\int_0^1(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(t_k,X_{t_k}^{\theta^0})\widehat{\mathrm{E}}^{\theta^0}\left[M_{i,2,k,n}|\widehat{\mathcal{F}}_{t_k}\right]d\ell\right|\\
&=\left|\varphi^i_{n\Delta_n}(\theta^0)u_i\sum_{k=0}^{n-1}\int_0^1\widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(t_k,X_{t_k}^{\theta^0})\right.\right.\\
&\qquad\left.\left.\cdot\int_{t_k}^{t_{k+1}}\left(b(\theta_i^{0+}(\ell),s,X_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta_i^{0+}(\ell)})\right)ds\left(\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}}-1\right)\right]d\ell\right|\\
&\leq\varphi^i_{n\Delta_n}(\theta^0)|u_i|\sum_{k=0}^{n-1}\int_0^1\left|\widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta^0}))^*(\sigma\sigma^*)^{-1}(t_k,X_{t_k}^{\theta^0})\right.\right.\\
&\qquad\left.\left.\cdot\int_{t_k}^{t_{k+1}}\left(b(\theta_i^{0+}(\ell),s,X_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t_k,X_{t_k}^{\theta_i^{0+}(\ell)})\right)ds\left(\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}}-1\right)\right]\right|d\ell\\
&\leq C\varphi^i_{n\Delta_n}(\theta^0)|u_i|\sum_{k=0}^{n-1}\int_0^1\sqrt{\Delta_n}\left(1+|X_{t_k}^{\theta^0}|\right)\left(\left|\int_{\theta_i^0+\ell\varphi^i_{n\Delta_n}(\theta^0)u_i}^{\theta_i^0}\left(\widehat{\mathrm{E}}^{\theta_i(0+)}_{t_k,X^{\theta^0}_{t_k}}\left[|V|^2\right]\right)^{\frac{1}{2}}d\theta_i\right|\right.\\
&\qquad+\left|\int_{\theta_{i+1}^0+\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}}^{\theta_{i+1}^0}\left(\widehat{\mathrm{E}}^{\theta_{i+1}(0+)}_{t_k,X^{\theta^0}_{t_k}}\left[|V|^2\right]\right)^{\frac{1}{2}}d\theta_{i+1}\right|\\
&\qquad\left.+\cdots+\left|\int_{\theta_m^0+\varphi^m_{n\Delta_n}(\theta^0)u_m}^{\theta_m^0}\left(\widehat{\mathrm{E}}^{\theta_m(0+)}_{t_k,X^{\theta^0}_{t_k}}\left[|V|^2\right]\right)^{\frac{1}{2}}d\theta_m\right|\right)d\ell\\
&\leq C\Delta_n^2\varphi^i_{n\Delta_n}(\theta^0)|u_i|\big(\varphi^i_{n\Delta_n}(\theta^0)|u_i|+\varphi^{i+1}_{n\Delta_n}(\theta^0)|u_{i+1}|+\cdots+\varphi^m_{n\Delta_n}(\theta^0)|u_m|\big)\sum_{k=0}^{n-1}\left(1+|X_{t_k}^{\theta^0}|^2\right)\\
&+C\Delta_n^{\frac{5}{2}}\varphi^i_{n\Delta_n}(\theta^0)|u_i|\big(\varphi^i_{n\Delta_n}(\theta^0)|u_i|+\varphi^{i+1}_{n\Delta_n}(\theta^0)|u_{i+1}|+\cdots+\varphi^m_{n\Delta_n}(\theta^0)|u_m|\big)\sum_{k=0}^{n-1}\left(1+|X_{t_k}^{\theta^0}|^3\right)\\
&\leq Cn\Delta_n^2\Big(|u_i|^2(\varphi^i_{n\Delta_n}(\theta^0))^2+|u_i||u_{i+1}|\varphi^i_{n\Delta_n}(\theta^0)\varphi^{i+1}_{n\Delta_n}(\theta^0)+\cdots+|u_i||u_m|\varphi^i_{n\Delta_n}(\theta^0)\varphi^m_{n\Delta_n}(\theta^0)\Big)
\end{aligned}$$

$$+ C|u_i|^2 n\Delta_n^2 \max_{k\in\{0,\dots,n-1\}} |\varphi^i_{t_k}(\theta^0)X^{\theta^0}_{t_k}|^2 + C|u_i||u_{i+1}|n\Delta_n^2 \max_{k\in\{0,\dots,n-1\}} \varphi^i_{t_k}(\theta^0)\varphi^{i+1}_{t_k}(\theta^0)|X^{\theta^0}_{t_k}|^2$$
$$+ \cdots + C|u_i||u_m|n\Delta_n^2 \max_{k\in\{0,\dots,n-1\}} \varphi^i_{t_k}(\theta^0)\varphi^{m}_{t_k}(\theta^0)|X^{\theta^0}_{t_k}|^2$$
$$+ Cn\Delta_n^{\frac{5}{2}}\Big(|u_i|^2(\varphi^i_{n\Delta_n}(\theta^0))^2 + |u_i||u_{i+1}|\varphi^i_{n\Delta_n}(\theta^0)\varphi^{i+1}_{n\Delta_n}(\theta^0) + \cdots + |u_i||u_m|\varphi^i_{n\Delta_n}(\theta^0)\varphi^{m}_{n\Delta_n}(\theta^0)\Big)$$
$$+ C|u_i|^2 n\Delta_n^{\frac{5}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^{-1} \max_{k\in\{0,\dots,n-1\}} |\varphi^i_{t_k}(\theta^0)X^{\theta^0}_{t_k}|^3$$
$$+ C|u_i||u_{i+1}|n\Delta_n^{\frac{5}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^{-1} \max_{k\in\{0,\dots,n-1\}} (\varphi^i_{t_k}(\theta^0))^2\varphi^{i+1}_{t_k}(\theta^0)|X^{\theta^0}_{t_k}|^3$$
$$+ \cdots + C|u_i||u_m|n\Delta_n^{\frac{5}{2}}(\varphi^i_{n\Delta_n}(\theta^0))^{-1} \max_{k\in\{0,\dots,n-1\}} (\varphi^i_{t_k}(\theta^0))^2\varphi^{m}_{t_k}(\theta^0)|X^{\theta^0}_{t_k}|^3,$$

which, by conditions **(A5)-(A6)**, converges to zero in $\widehat{\mathrm{P}}^{\theta^0}$-probability as $n\to\infty$. Here,

$$V := (\partial_{\theta_i} b(\theta_i^{0+}(\ell), t_k, X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k, X^{\theta^0}_{t_k}) \cdot \int_{t_k}^{t_{k+1}} \Big(b(\theta_i^{0+}(\ell), s, X_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta_i^{0+}(\ell)})\Big)ds.$$

and we have used the mean value theorem for vector-valued functions,

$$b(\theta_i^{0+}(\ell), s, X_s^{\theta_i^{0+}(\ell)}) - b(\theta_i^{0+}(\ell), t_k, X_{t_k}^{\theta_i^{0+}(\ell)}) = \left(\int_0^1 J_b(t_k + v(s-t_k), X_{t_k}^{\theta_i^{0+}(\ell)} + v(X_s^{\theta_i^{0+}(\ell)} - X_{t_k}^{\theta_i^{0+}(\ell)}))dv\right)\cdot \begin{pmatrix} s-t_k \\ X_s^{\theta_i^{0+}(\ell)} - X_{t_k}^{\theta_i^{0+}(\ell)}\end{pmatrix},$$

where the Jacobian matrix is given by

$$J_b(t_k + v(s-t_k), X_{t_k}^{\theta_i^{0+}(\ell)} + v(X_s^{\theta_i^{0+}(\ell)} - X_{t_k}^{\theta_i^{0+}(\ell)})) = \begin{pmatrix} \partial_t b_1 & \partial_{x_1} b_1 & \dots & \partial_{x_d} b_1 \\ \vdots & \vdots & \ddots & \vdots \\ \partial_t b_d & \partial_{x_1} b_d & \dots & \partial_{x_d} b_d \end{pmatrix} (\theta_i^{0+}(\ell), t_k + v(s-t_k), X_{t_k}^{\theta_i^{0+}(\ell)} + v(X_s^{\theta_i^{0+}(\ell)} - X_{t_k}^{\theta_i^{0+}(\ell)})).$$

Therefore, $\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}[\zeta_{i,2,k,n}|\widehat{\mathcal{F}}_{t_k}] \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} 0$ as $n\to\infty$.

Next, applying Jensen's inequality and Lemma 3.6, conditions **(A1)-(A2)**, the mean value theorem for vector-valued functions and Lemma 3.4 (i), we obtain that

$$\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[\zeta^2_{i,2,k,n}|\widehat{\mathcal{F}}_{t_k}\right]$$
$$= (\varphi^i_{n\Delta_n}(\theta^0)u_i)^2\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\left(\int_0^1(\partial_{\theta_i} b(\theta_i^{0+}(\ell), t_k, X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k, X^{\theta^0}_{t_k})M_{i,2,k,n}d\ell\right)^2\right]$$
$$\leq (\varphi^i_{n\Delta_n}(\theta^0)u_i)^2\sum_{k=0}^{n-1}\int_0^1\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\left|(\partial_{\theta_i} b(\theta_i^{0+}(\ell), t_k, X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k, X^{\theta^0}_{t_k})M_{i,2,k,n}\right|^2\right]d\ell$$
$$\leq 2(\varphi^i_{n\Delta_n}(\theta^0)u_i)^2\sum_{k=0}^{n-1}\int_0^1\left\{\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\left|(\partial_{\theta_i} b(\theta_i^{0+}(\ell), t_k, X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k, X^{\theta^0}_{t_k})M_{i,2,1,k,n}\right|^2\right]\right.$$

$$+\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[\left|(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k,X^{\theta^0}_{t_k})M_{i,2,2,k,n}\right|^2\Big|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\right]\Bigg\}d\ell$$

$$=2(\varphi^i_{n\Delta_n}(\theta^0)u_i)^2\sum_{k=0}^{n-1}\int_0^1\left\{\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[\left|(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k,X^{\theta^0}_{t_k})M_{i,2,1,k,n}\right|^2\right]\right.$$

$$\left.+\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[\left|(\partial_{\theta_i}b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k}))^*(\sigma\sigma^*)^{-1}(t_k,X^{\theta^0}_{t_k})M_{i,2,2,k,n}\right|^2\right]\right\}d\ell$$

$$\le C\Delta_n^3(\varphi^i_{n\Delta_n}(\theta^0)u_i)^2\sum_{k=0}^{n-1}\left(1+|X^{\theta^0}_{t_k}|^2\right)+C\Delta_n^4(\varphi^i_{n\Delta_n}(\theta^0)u_i)^2\sum_{k=0}^{n-1}\left(1+|X^{\theta^0}_{t_k}|^4\right)$$

$$\le C|u_i|^2n\Delta_n^3(\varphi^i_{n\Delta_n}(\theta^0))^2+C|u_i|^2n\Delta_n^3\max_{k\in\{0,\dots,n-1\}}|\varphi^i_{t_k}(\theta^0)X^{\theta^0}_{t_k}|^2$$

$$+C|u_i|^2n\Delta_n^4(\varphi^i_{n\Delta_n}(\theta^0))^2+C|u_i|^2n\Delta_n^4(\varphi^i_{n\Delta_n}(\theta^0))^{-2}\max_{k\in\{0,\dots,n-1\}}|\varphi^i_{t_k}(\theta^0)X^{\theta^0}_{t_k}|^4,$$

which, by conditions **(A5)**-**(A6)**, converges to zero in $\widehat{\mathrm{P}}^{\theta^0}$-probability as $n\to\infty$. Here

$$M_{i,2,1,k,n}=\int_{t_k}^{t_{k+1}}\left(b(\theta_i^{0+}(\ell),s,X^{\theta^0}_s)-b(\theta_i^{0+}(\ell),t_k,X^{\theta^0}_{t_k})\right)ds,$$

$$M_{i,2,2,k,n}=\int_{t_k}^{t_{k+1}}\left(b(\theta_i^{0+}(\ell),s,Y^{\theta_i^{0+}(\ell)}_s)-b(\theta_i^{0+}(\ell),t_k,Y^{\theta_i^{0+}(\ell)}_{t_k})\right)ds.$$

Thus, by Lemma 3.8 a), $\sum_{k=0}^{n-1}\zeta_{i,2,k,n}\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0$ for any $i\in\{1,\dots,m\}$. Thus, the result follows. □

Lemma 4.6. *Under conditions* **(A1)**-**(A3)** *and* **(A5)**-**(A6)**, *as $n\to\infty$,*

$$\sum_{k=0}^{n-1}\sum_{i=1}^m\frac{\varphi^i_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1\left\{Z^{5,\ell}_{i,k,n}-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[R_5^{\theta_i^{0+}(\ell),k}\Big|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\right\}d\ell\overset{\widehat{\mathrm{P}}^{\theta^0}}{\longrightarrow}0.$$

Proof. For any $i\in\{1,\dots,m\}$, we set

$$\zeta_{k,n}=\zeta_{i,k,n}:=\frac{\varphi^i_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1\left\{Z^{5,\ell}_{i,k,n}-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[R_5^{\theta_i^{0+}(\ell),k}\Big|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]\right\}d\ell.$$

Using Lemma 3.6, we get that

$$\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[\zeta_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\right]=\sum_{k=0}^{n-1}\frac{\varphi^i_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1\widehat{\mathrm{E}}^{\theta^0}\left[Z^{5,\ell}_{i,k,n}-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}\left[R_5^{\theta_i^{0+}(\ell),k}|Y^{\theta_i^{0+}(\ell)}_{t_{k+1}}=X^{\theta^0}_{t_{k+1}}\right]|\widehat{\mathcal{F}}_{t_k}\right]d\ell$$

$$=\sum_{k=0}^{n-1}\frac{\varphi^i_{n\Delta_n}(\theta^0)u_i}{\Delta_n}\int_0^1\left(\widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[Z^{5,\ell}_{i,k,n}\right]-\widetilde{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,X^{\theta^0}_{t_k}}[R_5^{\theta_i^{0+}(\ell),k}]\right)d\ell=0.$$

Next, proceeding as in the proof of Lemma 4.5, we obtain that

$$\sum_{k=0}^{n-1}\widehat{\mathrm{E}}^{\theta^0}\left[\zeta^2_{i,k,n}|\widehat{\mathcal{F}}_{t_k}\right]\le C\Delta_n^2(\varphi^i_{n\Delta_n}(\theta^0)u_i)^2\sum_{k=0}^{n-1}\left(1+|X^{\theta^0}_{t_k}|^2\right)$$

$$\le C|u_i|^2n\Delta_n^2(\varphi^i_{n\Delta_n}(\theta^0))^2+C|u_i|^2n\Delta_n^2\max_{k\in\{0,\dots,n-1\}}|\varphi^i_{t_k}(\theta^0)X^{\theta^0}_{t_k}|^2,$$

which, by conditions **(A5)-(A6)**, converges to zero in $\widehat{\mathrm{P}}^{\theta^0}$-probability as $n \to \infty$. Thus, by Lemma 3.8 a), we have shown that $\sum_{k=0}^{n-1} \zeta_{i,k,n} \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} 0$ for any $i \in \{1, \ldots, m\}$. Thus, the result follows. □

Lemma 4.7. *Under conditions* **(A1)**-**(A3)** *and* **(A5)-(A6)**, *as* $n \to \infty$,

$$-\sum_{k=0}^{n-1} H_{1,k,n} + \frac{1}{2}\sum_{k=0}^{n-1} H_{2,k,n} - \sum_{k=0}^{n-1}\sum_{i=1}^{m} (K_{i,k,n} + K_{i+1,k,n} + \cdots + K_{m,k,n})$$
$$+ \sum_{k=0}^{n-1}\sum_{i=1}^{m} \xi_{2,i,k,n} \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} 0.$$

Proof. We proceed as in the proof of Lemmas 4.4 4.5 and 4.6. □

Remark 4.8. *To obtain the results in Corollary 2.2, 2.3 and 2.4, the proof follows along the same lines as that of Theorem 2.1 except that condition* **(A5)** *is now replaced by condition* **(A5')** *or is removed, condition* **(A6)** *is now replaced by condition* **(A6')** *or* **(A6")** *or* **(A6"')**.

5. Examples

5.1. Homogeneous diffusions.

5.1.1. *Homogeneous ergodic diffusion processes.* Let $X^\theta = (X_t^\theta)_{t\geq 0}$ be the unique strong solution of the d-dimensional SDE

$$dX_t^\theta = b(\theta, X_t^\theta)dt + \sigma(X_t^\theta)dB_t. \tag{5.1}$$

This is a particular case of the model discussed in [7] where the unknown parameter appears only in the drift coefficient and when equation is homogeneous. We introduce the following ergodic assumption.

(A4') The process X^θ given by (5.1) is ergodic in the sense that there exists a unique probability measure $\pi_\theta(dx)$ such that as $T \to \infty$,

$$\frac{1}{T}\int_0^T g(X_t^\theta)dt \xrightarrow{\widehat{\mathrm{P}}^{\theta}} \int_{\mathbb{R}^d} g(x)\pi_\theta(dx),$$

for any π_θ-integrable function g.

Then, under condition **(A4')**, condition **(A4)** satisfies with $m\times m$ diagonal matrix $\varphi_T(\theta) = \mathrm{diag}(\frac{1}{\sqrt{T}}, \ldots, \frac{1}{\sqrt{T}})$ whose diagonal entries are chosen as $\varphi_T^1(\theta) = \cdots = \varphi_T^m(\theta) = \frac{1}{\sqrt{T}}$ and

$$\Gamma(\theta) = \int_{\mathbb{R}^d} (\nabla_\theta b(\theta, x))^* \, (\sigma^{-1}(x))^* \sigma^{-1}(x) \nabla_\theta b(\theta, x)\pi_\theta(dx).$$

Then, by ergodicity, as $n \to \infty$,

$$\frac{1}{n}\sum_{k=0}^{n-1} |X_{t_k}^{\theta^0}| \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} \int_{\mathbb{R}^d} |x|\pi_{\theta^0}(dx), \qquad \text{and} \qquad \frac{1}{n}\sum_{k=0}^{n-1} |X_{t_k}^{\theta^0}|^2 \xrightarrow{\widehat{\mathrm{P}}^{\theta^0}} \int_{\mathbb{R}^d} |x|^2\pi_{\theta^0}(dx).$$

In this case, condition **(A6)** is not required. Condition **(A5)** is reformulated as follows

(A5") For all $\theta \in \Theta$, $\int_{\mathbb{R}^d} |x|^2 \pi_\theta(dx) < \infty$.

As a consequence of Theorem 2.1, under conditions **(A1)-(A3)**, **(A4')** and **(A5")**, the LAN property holds at θ^0 with rate of convergence $\varphi_{n\Delta_n}(\theta^0) = \text{diag}(\frac{1}{\sqrt{n\Delta_n}}, \ldots, \frac{1}{\sqrt{n\Delta_n}})$ where $\varphi^1_{n\Delta_n}(\theta) = \cdots = \varphi^m_{n\Delta_n}(\theta) = \frac{1}{\sqrt{n\Delta_n}}$ and asymptotic Fisher information matrix

$$\Gamma(\theta^0) = \int_{\mathbb{R}^d} \left(\nabla_\theta b(\theta^0, x)\right)^* (\sigma^{-1}(x))^* \sigma^{-1}(x) \nabla_\theta b(\theta^0, x) \pi_{\theta^0}(dx).$$

That is, for all $u \in \mathbb{R}^m$, as $n \to \infty$,

$$\log \frac{d\mathrm{P}_n^{\theta^0 + \varphi_{n\Delta_n}(\theta^0)u}}{d\mathrm{P}_n^{\theta^0}} \left(X^{n,\theta^0}\right) \overset{\mathcal{L}(\widehat{\mathrm{P}}^{\theta^0})}{\longrightarrow} u^* \mathcal{N}(0, \Gamma(\theta^0)) - \frac{1}{2} u^* \Gamma(\theta^0) u.$$

5.1.2. *Homogeneous Ornstein-Uhlenbeck process.* Let $X^{a,b} = (X_t^{a,b})_{t \geq 0}$ be the unique strong solution of the one-dimensional SDE

$$dX_t^{a,b} = (b - aX_t^{a,b})dt + \sigma dB_t, \tag{5.2}$$

with given initial condition $X_0^{a,b} = x_0$, $\theta = (a, b) \in \mathbb{R}^2$ are unknown parameters and $\sigma > 0$. By Itô's formula, the solution process is given by

$$X_t^{a,b} = X_0^{a,b} e^{-at} + \frac{b}{a}\left(1 - e^{-at}\right) + \sigma \int_0^t e^{-a(t-s)} dB_s. \tag{5.3}$$

The observed Fisher information process at (a, b) based on the continuous observation $(X_t^{a,b})_{t \in [0,T]}$ is given by

$$\frac{1}{\sigma^2} \begin{pmatrix} \int_0^T (X_t^{a,b})^2 dt & -\int_0^T X_t^{a,b} dt \\ -\int_0^T X_t^{a,b} dt & T \end{pmatrix}.$$

Case 1: $a > 0$. The solution $X^{a,b}$ is ergodic with invariant Gaussian distribution $\mathcal{N}(\frac{b}{a}, \frac{\sigma^2}{2a})$ (see [13, Example 1.26]). That is,

$$\pi_{a,b}(dx) = f(a, b, x)dx = \sqrt{\frac{a}{\pi\sigma^2}} \exp\left\{-\frac{(ax - b)^2}{a\sigma^2}\right\} dx.$$

By [13, Example 1.35], as $T \to \infty$,

$$\begin{aligned} \frac{1}{T}\int_0^T (X_t^{a,b})^2 dt &\overset{\widehat{\mathrm{P}}^{a,b}}{\longrightarrow} \int_{\mathbb{R}} |x|^2 \pi_{a,b}(dx) = \frac{b^2}{a^2} + \frac{\sigma^2}{2a}, \\ \frac{1}{T}\int_0^T X_t^{a,b} dt &\overset{\widehat{\mathrm{P}}^{a,b}}{\longrightarrow} \int_{\mathbb{R}} x \pi_{a,b}(dx) = \frac{b}{a}. \end{aligned}$$

Thus, condition **(A4)** satisfies with $\varphi_T(a, b) = \text{diag}(\frac{1}{\sqrt{T}}, \frac{1}{\sqrt{T}})$ and

$$\Gamma(a, b) = \frac{1}{a^2\sigma^2} \begin{pmatrix} b^2 + \frac{a\sigma^2}{2} & -ab \\ -ab & a^2 \end{pmatrix}.$$

Moreover, condition **(A5")** holds. As a consequence of Theorem 2.1 (Subsection 5.1.1), the LAN property holds at $\theta^0 = (a_0, b_0)$ with rate of convergence $\varphi_{n\Delta_n}(a_0, b_0) = \text{diag}(\frac{1}{\sqrt{n\Delta_n}}, \frac{1}{\sqrt{n\Delta_n}})$

and asymptotic Fisher information matrix

$$\Gamma(a_0,b_0)=\frac{1}{a_0^2\sigma^2}\begin{pmatrix} b_0^2+\dfrac{a_0\sigma^2}{2} & -a_0b_0 \\ -a_0b_0 & a_0^2\end{pmatrix}.$$

Case 2: $a<0$. From (5.3), it can be checked that $e^{at}X_t^{a,b}-X_0^{a,b}-\frac{b}{a}(e^{at}-1)=\sigma\int_0^t e^{as}dB_s$, $t\geq 0$ is a square integrable martingale. Thus, the martingale convergence theorem implies that as $t\to\infty$,

$$e^{at}X_t^{a,b}\to X_0^{a,b}-\frac{b}{a}+Z^a,\quad \widehat{\mathrm{P}}^{a,b}\text{-a.s.},$$

where $Z^a:=\sigma\int\limits_0^\infty e^{as}dB_s$ has Gaussian law $\mathcal{N}(0,-\frac{\sigma^2}{2a})$. Then, the integral version of the Toeplitz lemma implies that as $t\to\infty$,

$$\begin{aligned}\frac{\int_0^t X_s^{a,b}ds}{\int_0^t e^{-as}ds}&\to X_0^{a,b}-\frac{b}{a}+Z^a,\quad \widehat{\mathrm{P}}^{a,b}\text{-a.s.}\\ \frac{\int_0^t (X_s^{a,b})^2ds}{\int_0^t e^{-2as}ds}&\to \left(X_0^{a,b}-\frac{b}{a}+Z^a\right)^2,\quad \widehat{\mathrm{P}}^{a,b}\text{-a.s.}\end{aligned}$$

which deduces that as $t\to\infty$,

$$\begin{aligned}&\frac{1}{\sqrt{t}}e^{at}\int_0^t X_s^{a,b}ds\to 0,\quad \widehat{\mathrm{P}}^{a,b}\text{-a.s.}\\ &e^{2at}\int_0^t (X_s^{a,b})^2ds\to -\frac{1}{2a}\left(X_0^{a,b}-\frac{b}{a}+Z^a\right)^2,\quad \widehat{\mathrm{P}}^{a,b}\text{-a.s.}\end{aligned}$$

Thus, condition **(A4)** satisfies with $\varphi_T(a,b)=\mathrm{diag}(e^{aT},\frac{1}{\sqrt{T}})$ and

$$\Gamma(a,b)=\begin{pmatrix}-\frac{1}{2a\sigma^2}(x_0-\frac{b}{a}+Z^a)^2 & 0\\ 0 & \frac{1}{\sigma^2}\end{pmatrix}.$$

Observe that $|\partial_a b(\theta,x)|=|x|$ and $|\partial_b b(\theta,x)|=1$, where $b(\theta,x)=b-ax$. Hence condition **(A6")** holds for the parameter a and condition **(A6"')** holds for the parameter b. On the other hand, $\widehat{\mathrm{E}}^{a,b}[\sup_{t\geq 0}|e^{at}X_t^{a,b}|^p]<\infty$ for any $p>0$. Thus, condition **(A5')** holds. As a consequence of Theorem 2.1 (Corollary 2.3 and 2.4), the LAMN property holds at (a_0,b_0) with rate of convergence $\varphi_{n\Delta_n}(a_0,b_0)=\mathrm{diag}(e^{a_0n\Delta_n},\frac{1}{\sqrt{n\Delta_n}})$ and asymptotic random Fisher information matrix

$$\Gamma(a_0,b_0)=\begin{pmatrix}-\frac{1}{2a_0\sigma^2}(x_0-\dfrac{b_0}{a_0}+Z^{a_0})^2 & 0\\ 0 & \frac{1}{\sigma^2}\end{pmatrix}.$$

5.1.3. *Two-dimensional Gaussian diffusion process.* Let $X^\theta=(X_1^\theta,X_2^\theta)^*=(X_t^\theta)_{t\geq 0}$ be the unique strong solution of the 2-dimensional SDE

$$dX_t^\theta=A(\theta)X_t^\theta dt+dB_t,\tag{5.4}$$

with $X_0^\theta = 0$, where

$$A(\theta) = \begin{pmatrix} -\theta_1 & -\theta_2 \\ \theta_2 & -\theta_1 \end{pmatrix},$$

$B = (B_t)_{t\geq 0}$ is a 2-dimensional Brownian motion and $\Theta = \mathbb{R}^2$ (see [17, Subsection 4.1]). By Itô's formula,

$$X_t^\theta = e^{A(\theta)t} \int_0^t e^{-A(\theta)s} dB_s,$$

where

$$e^{A(\theta)t} = e^{-\theta_1 t} \begin{pmatrix} \cos\theta_2 t & -\sin\theta_2 t \\ \sin\theta_2 t & \cos\theta_2 t \end{pmatrix}.$$

The observed Fisher information process at $\theta = (\theta_1, \theta_2)$ based on the continuous observation $(X_s^\theta)_{s\in[0,t]}$ is given by $\int_0^t |X_s^\theta|^2 ds\ I_2$.

Case 1: $\theta_1 < 0$. As $t \to \infty$,

$$\begin{aligned} e^{-A(\theta)t} X_t^\theta &\longrightarrow \sqrt{-\frac{1}{2\theta_1}} V(\theta), \quad \widehat{\mathrm{P}}^\theta\text{-a.s.}, \\ -\theta_1 e^{2\theta_1 t} |X_t^\theta|^2 &\longrightarrow \frac{1}{2} |V(\theta)|^2, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.}, \\ 2\theta_1^2 e^{2\theta_1 t} \int_0^t |X_s^\theta|^2 ds\ I_2 &\longrightarrow \frac{1}{2} |V(\theta)|^2 I_2, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.}, \end{aligned}$$

where $V(\theta) \sim \mathcal{N}(0, I_2)$. Thus, condition **(A4)** satisfies with $\varphi_t^1(\theta) = \varphi_t^2(\theta) = -\sqrt{2}\theta_1 e^{\theta_1 t}$ and $\Gamma(\theta) = \frac{1}{2}|V(\theta)|^2 I_2$. On the other hand, conditions **(A5')** and **(A6")** hold. As a consequence of Theorem 2.1 (Corollary 2.3), the LAMN property holds for the likelihood at $\theta^0 = (\theta_1^0, \theta_2^0)$ with rate of convergence $\varphi_{n\Delta_n}(\theta^0) = \mathrm{diag}(-\sqrt{2}\theta_1^0 e^{\theta_1^0 n\Delta_n}, -\sqrt{2}\theta_1^0 e^{\theta_1^0 n\Delta_n})$ and asymptotic random Fisher information matrix $\Gamma(\theta^0) = \frac{1}{2}|V(\theta^0)|^2 I_2$.

Case 2: $\theta_1 > 0$. By ergodicity, as $t \to \infty$,

$$\begin{aligned} \frac{1}{t} \int_0^t |X_s^\theta|^2 ds &\longrightarrow \lim_{s\to\infty} \mathrm{E}[|X_s^\theta|^2] = \int_{\mathbb{R}^2} |x|^2 \pi_\theta(dx) = \frac{1}{\theta_1}, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.}, \\ \frac{1}{t} \int_0^t |X_s^\theta|^2 ds\ I_2 &\longrightarrow \lim_{s\to\infty} \mathrm{E}[|X_s^\theta|^2]\ I_2 = \frac{1}{\theta_1}\ I_2, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.} \end{aligned}$$

Thus, condition **(A4)** satisfies with $\varphi_t^1(\theta) = \varphi_t^2(\theta) = \frac{1}{\sqrt{t}}$ and $\Gamma(\theta) = \frac{1}{\theta_1} I_2$. On the other hand, condition **(A5")** holds. As a consequence of Theorem 2.1 (Subsection 5.1.1), the LAN property holds for the likelihood at $\theta^0 = (\theta_1^0, \theta_2^0)$ with rate of convergence $\varphi_{n\Delta_n}(\theta^0) = \mathrm{diag}(\frac{1}{\sqrt{n\Delta_n}}, \frac{1}{\sqrt{n\Delta_n}})$ and asymptotic Fisher information matrix $\Gamma(\theta^0) = \dfrac{1}{\theta_1^0} I_2$.

5.1.4. *Null-recurrent diffusion process.* Let $X^\theta = (X_t^\theta)_{t\geq 0}$ be the unique strong solution of the one-dimensional SDE

$$dX_t^\theta = -\theta \frac{X_t^\theta}{1 + (X_t^\theta)^2} dt + \sigma dB_t, \tag{5.5}$$

where $X_0^\theta = x_0$ and $\sigma > 0$ (see [13, Subsection 3.5.1]). Observe that $b(\theta, x) = -\theta \frac{x}{1+x^2}$ which satisfies $|\partial_\theta b(\theta, x)| \leq \frac{1}{2}$. Hence, condition **(A5)** is not required.

The observed Fisher information process at θ based on the continuous observation $(X_t^\theta)_{t\in[0,T]}$ is given by $\int_0^T \frac{(X_t^\theta)^2}{\sigma^2(1+(X_t^\theta)^2)^2}dt$.

Case 1: $\theta > \frac{\sigma^2}{2}$. The process X^θ is ergodic with invariant density

$$f(\theta, x) = \frac{1}{G(\theta)(1+x^2)^{\theta/\sigma^2}} \quad \text{with} \quad G(\theta) = \int_{-\infty}^{\infty} \frac{dx}{(1+x^2)^{\theta/\sigma^2}}.$$

By ergodicity, as $T \to \infty$,

$$\frac{1}{T}\int_0^T \frac{(X_t^\theta)^2}{\sigma^2(1+(X_t^\theta)^2)^2}dt \xrightarrow{\widehat{\mathrm{P}}^\theta} \Gamma(\theta),$$

where

$$\Gamma(\theta) := \frac{1}{\sigma^2 G(\theta)}\int_{-\infty}^{\infty} \frac{x^2}{(1+x^2)^{2+\theta/\sigma^2}}dx.$$

Thus, condition **(A4)** satisfies with $\varphi_T(\theta) = \frac{1}{\sqrt{T}}$ and $\Gamma(\theta)$. As a consequence of Theorem 2.1 (Subsection 5.1.1), the LAN property holds for the likelihood at θ^0 with rate of convergence $\varphi_{n\Delta_n}(\theta^0) = \frac{1}{\sqrt{n\Delta_n}}$ and asymptotic Fisher information $\Gamma(\theta^0)$.

Case 2: $-\frac{\sigma^2}{2} < \theta < \frac{\sigma^2}{2}$. We set $\gamma(\theta) := \frac{1}{2} + \frac{\theta}{\sigma^2}$ and

$$K_*(B, \gamma(\theta)) = \frac{\Gamma(1+\gamma(\theta))}{2(\gamma(\theta)^2 B)^{\gamma(\theta)}\Gamma(1-\gamma(\theta))},$$

where $\Gamma(\cdot)$ is the Gamma function and

$$B = \frac{2}{\sigma^2}\left(1 + \frac{2\theta}{\sigma^2}\right)^{-\frac{4\theta}{\sigma^2+2\theta}}.$$

Let η be a random variable with stable distribution function having the Laplace transform $\mathrm{E}[e^{-p\eta}] = e^{-p^\gamma}$. As $T \to \infty$,

$$\frac{1}{T^{\gamma(\theta)}}\int_0^T \frac{(X_t^\theta)^2}{\sigma^2(1+(X_t^\theta)^2)^2}dt \xrightarrow{\widehat{\mathrm{P}}^\theta} \Gamma(\theta),$$

where

$$\Gamma(\theta) := K_*(B, \gamma(\theta))\frac{2}{\sigma^2}\int_{-\infty}^{\infty} \frac{x^2}{(1+x^2)^{2+\theta/\sigma^2}}dx\, \eta^{-\gamma(\theta)}.$$

Thus, condition **(A4)** satisfies with $\varphi_T(\theta) = T^{-\frac{\gamma(\theta)}{2}}$ and $\Gamma(\theta)$. Condition **(A6'')** writes as $n\Delta_n^{\frac{2(\sigma^2-\theta)}{\sigma^2-2\theta}} \to 0$ as $n \to \infty$. As a consequence of Theorem 2.1 (Corollary 2.4), under condition $n\Delta_n^{\frac{2(\sigma^2-\theta^0)}{\sigma^2-2\theta^0}} \to 0$, the LAMN property holds for the likelihood at θ^0 with rate of convergence $\varphi_{n\Delta_n}(\theta^0) = (n\Delta_n)^{-\frac{\gamma(\theta^0)}{2}}$ and asymptotic random Fisher information $\Gamma(\theta^0)$.

5.1.5. *Exponential growth process.* Let $X^\theta = (X_t^\theta)_{t\geq 0}$ be the unique strong solution of the one-dimensional SDE

$$dX_t^\theta = \theta a(X_t^\theta)dt + dB_t, \tag{5.6}$$

with given initial condition $X_0^\theta = x_0$. The unknown parameter θ is positive. For some constant $c > 0$, the known trend coefficient admits the representation

$$a(x) = cx + r(x), \quad x \in \mathbb{R},$$

such that the function r satisfies the following Lipschitz and growth conditions with appropriate constants $K \geq 0$, $L \geq 0$ and $\gamma \in [0,1)$. That is, for all $(x,y) \in \mathbb{R}^2$,

$$|r(x) - r(y)| \leq L\,|x-y|\,,$$
$$|r(x)| \leq K\left(1 + |x|^\gamma\right).$$

See Dietz and Kutoyants [5]. When taking large value, X_t^θ behaves like an Ornstein-Uhlenbeck process. The observed Fisher information process at θ based on the continuous observation $(X_t^\theta)_{t\in[0,T]}$ is given by $\int_0^T a^2(X_t^\theta)dt$.

By [5, Lemma 2.1 and Corollary 2.2], as $t \to \infty$ and $T \to \infty$,

$$e^{-\theta ct}X_t^\theta \longrightarrow X_0^\theta + \xi_\infty^\theta + \rho_\infty^\theta, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.},$$
$$e^{-2\theta cT}\int_0^T a^2(X_t^\theta)dt \longrightarrow \frac{c}{2\theta}\left(X_0^\theta + \xi_\infty^\theta + \rho_\infty^\theta\right)^2, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.},$$

where $\xi_\infty^\theta = \int_0^\infty e^{-\theta cs}dB_s$ and $\rho_\infty^\theta = \int_0^\infty e^{-\theta cs}\theta r(X_s^\theta)ds$. Moreover, $\widehat{\mathrm{E}}^\theta[\sup_{t\geq 0}|e^{-\theta ct}X_t^\theta|^p] < \infty$ for $p \geq 1$. Thus, condition **(A4)** satisfies with $\varphi_T(\theta) = e^{-\theta cT}$ and $\Gamma(\theta) = \frac{c}{2\theta}(X_0^\theta + \xi_\infty^\theta + \rho_\infty^\theta)^2$. Moreover, conditions **(A5')** and **(A6")** hold. As a consequence of Theorem 2.1 (Corollary 2.3), the LAMN property holds for the likelihood at θ^0 with rate of convergence $\varphi_{n\Delta_n}(\theta^0) = e^{-\theta^0 cn\Delta_n}$ and asymptotic random Fisher information

$$\Gamma(\theta^0) = \frac{c}{2\theta^0}\left(X_0^{\theta^0} + \xi_\infty^{\theta^0} + \rho_\infty^{\theta^0}\right)^2.$$

5.2. **Inhomogeneous diffusions.**

5.2.1. *Inhomogeneous Ornstein-Uhlenbeck process.* Let $X^\theta = (X_t^\theta)_{t\geq 0}$ be the unique strong solution of the one-dimensional SDE

$$dX_t^\theta = -\theta A(t)X_t^\theta dt + dB_t, \tag{5.7}$$

where $X_0^\theta = 0$, $A : \mathbb{R}_+ \to \mathbb{R}$ is measurable with $\int_0^t A^2(s)ds < \infty$ for every t (see [17, Subsection 4.2]). By Itô's formula, $X_t^\theta = f(\theta,t)\int_0^t f(\theta,s)^{-1}dB_s$ where $f(\theta,t) = \exp\{-\theta\int_0^t A(s)ds\}$.

The observed Fisher information process at θ based on the continuous observation $(X_s^\theta)_{s\in[0,t]}$ is given by $\int_0^t A^2(s)(X_s^\theta)^2ds$.

Case 1: Consider the set of explosive parameters

$$\Theta_0 := \left\{\theta \in \mathbb{R} : -\theta\int_0^t A(s)ds \to \infty \ \text{ as } \ t\to\infty \ \text{ and } \ \int_0^\infty f(\theta,t)^{-2}dt < \infty\right\}.$$

For any $\theta \in \Theta_0$, we have $\sup_{t\geq 0} f(\theta,t)^{-2}\widehat{\mathrm{E}}^\theta[(X_t^\theta)^2] = \int_0^\infty f(\theta,t)^{-2}dt < \infty$ and as $t \to \infty$,

$$f(\theta,t)^{-1}X_t^\theta \longrightarrow \left(\int_0^\infty f(\theta,s)^{-2}ds\right)^{\frac{1}{2}} V(\theta), \quad \widehat{\mathrm{P}}^\theta\text{-a.s.},$$

$$\varphi_t(\theta)^2\int_0^t A^2(s)(X_s^\theta)^2 ds \longrightarrow V(\theta)^2, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.},$$

where $V(\theta) \sim \mathcal{N}(0,1)$ and

$$\varphi_t(\theta) = \left(\int_0^\infty f(\theta,s)^{-2}ds\int_0^t A^2(s)f(\theta,s)^2ds\right)^{-\frac{1}{2}}.$$

Thus, condition **(A4)** satisfies with $\varphi_t(\theta)$ and $\Gamma(\theta) = V(\theta)^2$. On the other hand, condition **(A5)** holds. As a consequence of Theorem 2.1, under condition **(A6)**, the LAMN property holds for the likelihood at $\theta^0 \in \Theta_0$ with rate of convergence $\varphi_{n\Delta_n}(\theta^0)$ and asymptotic random Fisher information $\Gamma(\theta^0) = V(\theta^0)^2$.

Case 2: Consider the set of parameters

$$\Theta_1 := \Bigg\{\theta \in \mathbb{R} : \int_0^t A^2(s)f(\theta,s)^2\int_0^s f(\theta,u)^{-2}duds \to \infty \text{ as } t\to\infty$$
$$\text{and } A^2(t)f(\theta,t)^4\left(\int_0^t f(\theta,s)^{-2}ds\right)^2 = o\left(\int_0^t A^2(s)f(\theta,s)^2\int_0^s f(\theta,u)^{-2}duds\right)\Bigg\},$$

where assume that A is continuous.

For any $\theta \in \Theta_1$, as $t\to\infty$,

$$\varphi_t(\theta)^2\int_0^t A^2(s)(X_s^\theta)^2ds \longrightarrow 1, \quad \text{in } L^2(\widehat{\mathrm{P}}^\theta),$$

where

$$\varphi_t(\theta) = \left(\int_0^t A^2(s)f(\theta,s)^2\int_0^s f(\theta,u)^{-2}duds\right)^{-\frac{1}{2}}.$$

Thus, condition **(A4)** satisfies with $\varphi_t(\theta)$ and $\Gamma(\theta) = 1$. On the other hand, condition **(A5)** holds. As a consequence of Theorem 2.1, under condition **(A6)**, the LAN property holds for the likelihood at $\theta^0 \in \Theta_1$ with rate of convergence $\varphi_{n\Delta_n}(\theta^0)$ and asymptotic Fisher information $\Gamma(\theta^0) = 1$.

When $A(t) = 1$, X^θ becomes the classical homogeneous Ornstein-Uhlenbeck process which has been addressed in Subsection 5.1.2.

When $A(t) = \frac{1}{1+t}$, then $\Theta_0 = (-\infty, -\frac{1}{2})$ and $\Theta_1 = (-\frac{1}{2}, \infty)$. For any $\theta \in \Theta_0$, we choose $\varphi_T(\theta) = -(2\theta+1)T^{\theta+\frac{1}{2}}$ and for any $\theta \in \Theta_1$, we choose $\varphi_T(\theta) = \sqrt{\frac{2\theta+1}{\log(1+T)}}$.

5.2.2. *A special inhomogeneous diffusion process.* Let $X^\theta = (X_t^\theta)_{t\geq 0}$ be the unique strong solution of the one-dimensional SDE which is a special case of Hull-White model

$$dX_t^\theta = \theta b(t)X_t^\theta dt + \sigma(t)dB_t, \tag{5.8}$$

with given initial condition $X_0^\theta = 0$, where $b : \mathbb{R}_+ \to \mathbb{R}$ and $\sigma : \mathbb{R}_+ \to (0,\infty)$ are known Borel-measurable functions. Here, $\theta \in \mathbb{R}$ is an unknown parameter. See Barczy et Pap [2].

When $\sigma(t) = \sigma > 0$, X^θ becomes the inhomogeneous Ornstein-Uhlenbeck process which has been considered in Subsection 5.2.1.

The SDE (5.8) has a unique strong solution given by

$$X_t^\theta = \int_0^t \sigma(s) \exp\left\{\theta \int_s^t b(u)du\right\} dB_s.$$

The observed Fisher information process at θ based on the continuous observation $(X_s^\theta)_{s\in[0,t]}$ is given by $\int_0^t \frac{b^2(s)(X_s^\theta)^2}{\sigma^2(s)} ds$. The expected Fisher information at θ and t based on the continuous observation $(X_s^\theta)_{s\in[0,t]}$ is given by

$$I_{X^\theta}(t) = \int_0^t \frac{b^2(s)}{\sigma^2(s)} \mathrm{E}\left[(X_s^\theta)^2\right] ds,$$

where

$$\mathrm{E}\left[(X_s^\theta)^2\right] = \int_0^s \sigma^2(u) \exp\left\{2\theta \int_u^s b(v)dv\right\} du.$$

Case 1: Consider the set of parameters

$$\Theta_0 := \left\{\theta \in \mathbb{R} : \lim_{t\to\infty} I_{X^\theta}(t) = \infty \quad \text{and} \quad \lim_{t\to\infty} \int_0^t \sigma^2(s) \exp\left\{-2\theta \int_0^s b(v)dv\right\} ds < \infty\right\}.$$

Then, for any $\theta \in \Theta_0$, as $t \to \infty$,

$$\frac{1}{I_{X^\theta}(t)} \int_0^t \frac{b^2(s)(X_s^\theta)^2}{\sigma^2(s)} ds \longrightarrow \xi^2, \quad \widehat{\mathrm{P}}^\theta\text{-a.s.},$$

where $\xi \sim \mathcal{N}(0,1)$ (see the proof of [2, Theorem 7]).

Thus, condition **(A4)** satisfies with $\varphi_t(\theta) = I_{X^\theta}(t)^{-\frac{1}{2}}$ and $\Gamma(\theta) = \xi^2$. On the other hand, condition **(A5)** holds. As a consequence of Theorem 2.1, under condition **(A6)**, the LAMN property holds for the likelihood at θ^0 with rate of convergence $\varphi_{n\Delta_n}(\theta^0) = I_{X^{\theta^0}}(n\Delta_n)^{-\frac{1}{2}}$ and asymptotic random Fisher information $\Gamma(\theta^0) = \xi^2$.

Case 2: Consider the set of parameters

$$\Theta_1 := \left\{\theta \in \mathbb{R} : \lim_{t\to\infty} I_{X^\theta}(t) = \infty \text{ and } \lim_{t\to\infty} \frac{1}{\sqrt{I_{X^\theta}(t)}} \frac{b(t)}{\sigma^2(t)} \int_0^t \sigma^2(s) \exp\left\{2\theta \int_s^t b(v)dv\right\} ds = 0\right\}.$$

Then, for any $\theta \in \Theta_1$, as $t \to \infty$, (see [2, Theorem 10])

$$\frac{1}{I_{X^\theta}(t)} \int_0^t \frac{b^2(s)(X_s^\theta)^2}{\sigma^2(s)} ds \longrightarrow 1, \quad \text{in } L^2(\widehat{\mathrm{P}}^\theta).$$

Thus, condition **(A4)** satisfies with $\varphi_t(\theta) = I_{X^\theta}(t)^{-\frac{1}{2}}$ and $\Gamma(\theta) = 1$. On the other hand, condition **(A5)** holds. As a consequence of Theorem 2.1, under condition **(A6)**, the LAN property holds for the likelihood at θ^0 with rate of convergence $\varphi_{n\Delta_n}(\theta^0) = I_{X^{\theta^0}}(n\Delta_n)^{-\frac{1}{2}}$ and asymptotic Fisher information $\Gamma(\theta^0) = 1$.

6. Appendix

6.1. **Proof of Lemma 3.1.**

Proof. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a continuously differentiable function with compact support. Fix $t \in [t_k, t_{k+1}]$. The chain rule of the Malliavin calculus gives $(D_t(f(Y^\theta_{t_{k+1}}(t_k,x))))^* = (\nabla f(Y^\theta_{t_{k+1}}(t_k,x)))^* D_t Y^\theta_{t_{k+1}}(t_k,x)$. Since the matrix $D_t Y^\theta_{t_{k+1}}(t_k,x)$ is invertible a.s., we have $(\nabla f(Y^\theta_{t_{k+1}}(t_k,x)))^* = (D_t(f(Y^\theta_{t_{k+1}}(t_k,x))))^* U^\theta_t(t_k,x)$, where $U^\theta_t(t_k,x) = (D_t Y^\theta_{t_{k+1}}(t_k,x))^{-1}$.

Then, using the integration by parts formula of the Malliavin calculus on $[t_k, t_{k+1}]$, we get that for any $i \in \{1,\dots,m\}$,

$$\begin{aligned}
\partial_{\theta_i}\widetilde{\mathrm{E}}\left[f(Y^\theta_{t_{k+1}}(t_k,x))\right] &= \widetilde{\mathrm{E}}\left[(\nabla f(Y^\theta_{t_{k+1}}(t_k,x)))^* \partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)\right] \\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}\left[\int_{t_k}^{t_{k+1}} (\nabla f(Y^\theta_{t_{k+1}}(t_k,x)))^* \partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)dt\right] \\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}\left[\int_{t_k}^{t_{k+1}} (D_t(f(Y^\theta_{t_{k+1}}(t_k,x))))^* U^\theta_t(t_k,x)\, \partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)dt\right] \\
&= \frac{1}{\Delta_n}\widetilde{\mathrm{E}}\left[f(Y^\theta_{t_{k+1}}(t_k,x))\delta\left(U^\theta(t_k,x)\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)\right)\right].
\end{aligned}$$

Observe that by (3.7), the family $((\nabla f(Y^\theta_{t_{k+1}}(t_k,x)))^* \partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x), \theta \in \Theta)$ is uniformly integrable. This justifies that we can interchange ∂_{θ_i} and $\widetilde{\mathrm{E}}$. Note that here $\delta(V) \equiv \delta(V\mathbf{1}_{[t_k,t_{k+1}]}(\cdot))$ for any $V \in \mathrm{Dom}\ \delta$. On the other hand, using the stochastic flow property, we have that

$$\partial_{\theta_i}\widetilde{\mathrm{E}}\left[f(Y^\theta_{t_{k+1}}(t_k,x))\right] = \int_{\mathbb{R}^d} f(y)\partial_{\theta_i} p^\theta(t_k,t_{k+1},x,y)dy,$$

and

$$\begin{aligned}
&\widetilde{\mathrm{E}}\left[f(Y^\theta_{t_{k+1}}(t_k,x))\delta\left(U^\theta(t_k,x)\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)\right)\right] \\
&= \widetilde{\mathrm{E}}\left[f(Y^\theta_{t_{k+1}})\delta\left(U^\theta(t_k,x)\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)\right)\Big| Y^\theta_{t_k} = x\right] \\
&= \int_{\mathbb{R}^d} f(y)\widetilde{\mathrm{E}}\left[\delta\left(U^\theta(t_k,x)\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)\right)\Big| Y^\theta_{t_k} = x, Y^\theta_{t_{k+1}} = y\right] p^\theta(t_k,t_{k+1},x,y)dy,
\end{aligned}$$

which finishes the desired proof. □

6.2. **Proof of Lemma 3.2.**

Proof. From (3.4) and Itô's formula,

$$\begin{aligned}
(\nabla_x Y^\theta_t(t_k,x))^{-1} &= \mathrm{I}_d - \int_{t_k}^t (\nabla_x Y^\theta_s(t_k,x))^{-1}\left(\nabla_x b(\theta,s,Y^\theta_s(t_k,x)) - \sum_{j=1}^d (\nabla_x \sigma_j(s,Y^\theta_s(t_k,x)))^2\right) ds \\
&\quad - \sum_{j=1}^d \int_{t_k}^t (\nabla_x Y^\theta_s(t_k,x))^{-1}\nabla_x\sigma_j(s,Y^\theta_s(t_k,x))dW^j_s,
\end{aligned}$$

which, together with (3.5) and Itô's formula again, implies that

$$(\nabla_x Y^\theta_{t_{k+1}}(t_k,x))^{-1}\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x) = \int_{t_k}^{t_{k+1}} (\nabla_x Y^\theta_s(t_k,x))^{-1}\partial_{\theta_i} b(\theta,s,Y^\theta_s(t_k,x))ds. \tag{6.1}$$

Then, using the product rule [19, (1.48)] and (6.1), we obtain that

$$\begin{aligned}
&\delta\left(U^\theta(t_k,x)\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)\right)\\
&= \delta\left(\sigma^{-1}(\cdot, Y^\theta_\cdot(t_k,x))\nabla_x Y^\theta_\cdot(t_k,x)(\nabla_x Y^\theta_{t_{k+1}}(t_k,x))^{-1}\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)\right)\\
&= (\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x))^*((\nabla_x Y^\theta_{t_{k+1}}(t_k,x))^{-1})^* \int_{t_k}^{t_{k+1}} (\nabla_x Y^\theta_s(t_k,x))^*(\sigma^{-1}(s,Y^\theta_s(t_k,x)))^* dW_s\\
&\quad - \int_{t_k}^{t_{k+1}} \operatorname{tr}\left(D_s\left((\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x))^*((\nabla_x Y^\theta_{t_{k+1}}(t_k,x))^{-1})^*\right)\sigma^{-1}(s,Y^\theta_s(t_k,x))\nabla_x Y^\theta_s(t_k,x)\right)ds\\
&= \int_{t_k}^{t_{k+1}} ((\nabla_x Y^\theta_s(t_k,x))^{-1}\partial_{\theta_i} b(\theta,s,Y^\theta_s(t_k,x)))^* ds \int_{t_k}^{t_{k+1}} (\nabla_x Y^\theta_s(t_k,x))^*(\sigma^{-1}(s,Y^\theta_s(t_k,x)))^* dW_s\\
&- \int_{t_k}^{t_{k+1}}\int_s^{t_{k+1}} \operatorname{tr}\left(D_s\left(((\nabla_x Y^\theta_u(t_k,x))^{-1}\partial_{\theta_i} b(\theta,u,Y^\theta_u(t_k,x)))^*\right)\sigma^{-1}(s,Y^\theta_s(t_k,x))\nabla_x Y^\theta_s(t_k,x)\right)duds.
\end{aligned}$$

We next add and subtract the matrix $((\nabla_x Y^\theta_{t_k}(t_k,x))^{-1}\partial_{\theta_i} b(\theta,t_k,Y^\theta_{t_k}(t_k,x)))^*$ in the first integral and the matrix $(\nabla_x Y^\theta_{t_k}(t_k,x))^*(\sigma^{-1}(t_k,Y^\theta_{t_k}(t_k,x)))^*$ in the second integral. This, together with the fact that $Y^\theta_{t_k}(t_k,x) = Y^\theta_{t_k} = x$, yields

$$\delta\left(U^\theta(t_k,x)\partial_{\theta_i} Y^\theta_{t_{k+1}}(t_k,x)\right) = \Delta_n(\sigma^{-1}(t_k,x)\partial_{\theta_i} b(\theta,t_k,x))^*(W_{t_{k+1}} - W_{t_k}) - R_1^{\theta,k} + R_2^{\theta,k} + R_3^{\theta,k}. \tag{6.2}$$

On the other hand, by equation (3.3) we have that

$$\begin{aligned}
W_{t_{k+1}} - W_{t_k} = \sigma^{-1}(t_k,Y^\theta_{t_k})\Bigg(&Y^\theta_{t_{k+1}} - Y^\theta_{t_k} - b(\theta,t_k,Y^\theta_{t_k})\Delta_n - \int_{t_k}^{t_{k+1}}\left(b(\theta,s,Y^\theta_s) - b(\theta,t_k,Y^\theta_{t_k})\right)ds\\
&- \int_{t_k}^{t_{k+1}}\left(\sigma(s,Y^\theta_s) - \sigma(t_k,Y^\theta_{t_k})\right)dW_s\Bigg).
\end{aligned}$$

This, together with (6.2), conclude the desired result. □

6.3. **Proof of Lemma 3.6.**

Proof. For simplicity, we set $g(y) = g(X^{\theta^0}_{t_k}, y) := \widetilde{\mathrm{E}}^\theta_{t_k, X^{\theta^0}_{t_k}}[V \,|\, Y^\theta_{t_{k+1}} = y]$ for all $y \in \mathbb{R}^d$. Then, applying Girsanov's theorem, we obtain that

$$\begin{aligned}
&\widehat{\mathrm{E}}^{\theta^0}\left[\widetilde{\mathrm{E}}^\theta_{t_k,X^{\theta^0}_{t_k}}\left[V\,\middle|\,Y^\theta_{t_{k+1}} = X^{\theta^0}_{t_{k+1}}\right]\,\middle|\,\widehat{\mathcal{F}}_{t_k}\right] = \widehat{\mathrm{E}}^{\theta^0}\left[g(X^{\theta^0}_{t_{k+1}})\,\middle|\,X^{\theta^0}_{t_k}\right] = \widehat{\mathrm{E}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}\left[g(X^{\theta^0}_{t_{k+1}})\right]\\
&= \widehat{\mathrm{E}}^\theta_{t_k,X^{\theta^0}_{t_k}}\left[g(X^\theta_{t_{k+1}})\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^\theta_{t_k,X^{\theta^0}_{t_k}}}\left((X^\theta_t)_{t\in[t_k,t_{k+1}]}\right)\right]
\end{aligned}$$

$$
\begin{aligned}
&= \widehat{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[\widehat{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[g(X^{\theta}_{t_{k+1}}) \frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta}_{t_k, X^{\theta^0}_{t_k}}} \Big((X^{\theta}_t)_{t\in[t_k,t_{k+1}]} \Big) \Big| X^{\theta}_{t_{k+1}} \right] \right] \\
&= \widehat{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[g(X^{\theta}_{t_{k+1}}) \widehat{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta}_{t_k, X^{\theta^0}_{t_k}}} \Big((X^{\theta}_t)_{t\in[t_k,t_{k+1}]} \Big) \Big| X^{\theta}_{t_{k+1}} \right] \right] \\
&= \int_{\mathbb{R}^d} g(y) \widehat{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta}_{t_k, X^{\theta^0}_{t_k}}} ((X^{\theta}_t)_{t\in[t_k,t_{k+1}]}) \Big| X^{\theta}_{t_{k+1}} = y \right] p^{\theta}(t_k, t_{k+1}, X^{\theta^0}_{t_k}, y) dy \\
&= \int_{\mathbb{R}^d} \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[V \big| Y^{\theta}_{t_{k+1}} = y \right] \widehat{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta}_{t_k, X^{\theta^0}_{t_k}}} \Big((X^{\theta}_t)_{t\in[t_k,t_{k+1}]} \Big) \Big| X^{\theta}_{t_{k+1}} = y \right] p^{\theta}(t_k, t_{k+1}, X^{\theta^0}_{t_k}, y) dy \\
&= \int_{\mathbb{R}^d} \mathrm{E}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[V \frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta}_{t_k, X^{\theta^0}_{t_k}}} \Big((X^{\theta}_t)_{t\in[t_k,t_{k+1}]} \Big) \Big| X^{\theta}_{t_{k+1}} = y, Y^{\theta}_{t_{k+1}} = y \right] p^{\theta}(t_k, t_{k+1}, X^{\theta^0}_{t_k}, y) dy \\
&= \widehat{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[\mathrm{E}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[V \frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta}_{t_k, X^{\theta^0}_{t_k}}} ((X^{\theta}_t)_{t\in[t_k,t_{k+1}]}) \big| X^{\theta}_{t_{k+1}}, Y^{\theta}_{t_{k+1}} = X^{\theta}_{t_{k+1}} \right] \right] \\
&= \mathrm{E}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[\mathrm{E}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[V \frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta}_{t_k, X^{\theta^0}_{t_k}}} ((X^{\theta}_t)_{t\in[t_k,t_{k+1}]}) \big| X^{\theta}_{t_{k+1}}, Y^{\theta}_{t_{k+1}} = X^{\theta}_{t_{k+1}} \right] \right] \\
&= \mathrm{E}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[V \frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta}_{t_k, X^{\theta^0}_{t_k}}} ((X^{\theta}_t)_{t\in[t_k,t_{k+1}]}) \right] = \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} [V] \, \widehat{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta}_{t_k, X^{\theta^0}_{t_k}}} \Big((X^{\theta}_t)_{t\in[t_k,t_{k+1}]} \Big) \right] \\
&= \widetilde{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} [V],
\end{aligned}
$$

where we have used the fact that, by definition of $\mathrm{E}^{\theta}_{t_k,x}$, for any $\widehat{\mathcal{F}}_{t_{k+1}}$-measurable random variable V_1 and $\widetilde{\mathcal{F}}_{t_{k+1}}$-measurable random variable V_2,

$$
\widehat{\mathrm{E}}^{\theta}_{t_k,x} \left[V_1 | X^{\theta}_{t_{k+1}} = y \right] \widetilde{\mathrm{E}}^{\theta}_{t_k,x} \left[V_2 | Y^{\theta}_{t_{k+1}} = y \right] = \mathrm{E}^{\theta}_{t_k,x} \left[V_1 V_2 | X^{\theta}_{t_{k+1}} = y, Y^{\theta}_{t_{k+1}} = y \right],
$$

and $\widehat{\mathrm{E}}^{\theta}_{t_k, X^{\theta^0}_{t_k}} \left[\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta}_{t_k, X^{\theta^0}_{t_k}}} ((X^{\theta}_t)_{t\in[t_k,t_{k+1}]}) \right] = 1$ together with the independence between V and $\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k, X^{\theta^0}_{t_k}}}{d\widehat{\mathrm{P}}^{\theta}_{t_k, X^{\theta^0}_{t_k}}} ((X^{\theta}_t)_{t\in[t_k,t_{k+1}]})$ w.r.t. P. Thus, the result follows. □

6.4. **Proof of Lemma 3.7.**

Proof. Using (3.10), we have that

$$\begin{aligned}
&\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}-1=\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,x}-d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\\
&=\frac{(d\widehat{\mathrm{P}}^{\theta_{i+1}^{0+}}_{t_k,x}-d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x})+(d\widehat{\mathrm{P}}^{\theta_{i+2}^{0+}}_{t_k,x}-d\widehat{\mathrm{P}}^{\theta_{i+1}^{0+}}_{t_k,x})+\cdots+(d\widehat{\mathrm{P}}^{\theta_m^{0+}}_{t_k,x}-d\widehat{\mathrm{P}}^{\theta_{m-1}^{0+}}_{t_k,x})+(d\widehat{\mathrm{P}}^{\theta^0}_{t_k,x}-d\widehat{\mathrm{P}}^{\theta_m^{0+}}_{t_k,x})}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\\
&=\int_{\theta_i^0+\ell\varphi^i_{n\Delta_n}(\theta^0)u_i}^{\theta_i^0}\partial_{\theta_i}\left(\frac{d\widehat{\mathrm{P}}^{\theta_i(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\right)d\theta_i+\int_{\theta_{i+1}^0+\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}}^{\theta_{i+1}^0}\partial_{\theta_{i+1}}\left(\frac{d\widehat{\mathrm{P}}^{\theta_{i+1}(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\right)d\theta_{i+1}\\
&\qquad+\cdots+\int_{\theta_{m-1}^0+\varphi^{m-1}_{n\Delta_n}(\theta^0)u_{m-1}}^{\theta_{m-1}^0}\partial_{\theta_{m-1}}\left(\frac{d\widehat{\mathrm{P}}^{\theta_{m-1}(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\right)d\theta_{m-1}\\
&\qquad+\int_{\theta_m^0+\varphi^m_{n\Delta_n}(\theta^0)u_m}^{\theta_m^0}\partial_{\theta_m}\left(\frac{d\widehat{\mathrm{P}}^{\theta_m(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\right)d\theta_m\\
&=\int_{\theta_i^0+\ell\varphi^i_{n\Delta_n}(\theta^0)u_i}^{\theta_i^0}\int_{t_k}^{t_{k+1}}(\partial_{\theta_i}b(\theta_i(0+),t,X_t^{\theta_i^{0+}(\ell)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_i^{0+}(\ell)})\\
&\cdot\left(dB_t-\sigma^{-1}(t,X_t^{\theta_i^{0+}(\ell)})\big(b(\theta_i(0+),t,X_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t,X_t^{\theta_i^{0+}(\ell)})\big)dt\right)\frac{d\widehat{\mathrm{P}}^{\theta_i(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}d\theta_i\\
&+\int_{\theta_{i+1}^0+\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}}^{\theta_{i+1}^0}\int_{t_k}^{t_{k+1}}(\partial_{\theta_{i+1}}b(\theta_{i+1}(0+),t,X_t^{\theta_i^{0+}(\ell)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_i^{0+}(\ell)})\\
&\cdot\left(dB_t-\sigma^{-1}(t,X_t^{\theta_i^{0+}(\ell)})\big(b(\theta_{i+1}(0+),t,X_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t,X_t^{\theta_i^{0+}(\ell)})\big)dt\right)\frac{d\widehat{\mathrm{P}}^{\theta_{i+1}(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}d\theta_{i+1}\\
&+\cdots+\int_{\theta_m^0+\varphi^m_{n\Delta_n}(\theta^0)u_m}^{\theta_m^0}\int_{t_k}^{t_{k+1}}(\partial_{\theta_m}b(\theta_m(0+),t,X_t^{\theta_i^{0+}(\ell)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_i^{0+}(\ell)})\\
&\cdot\left(dB_t-\sigma^{-1}(t,X_t^{\theta_i^{0+}(\ell)})\big(b(\theta_m(0+),t,X_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t,X_t^{\theta_i^{0+}(\ell)})\big)dt\right)\frac{d\widehat{\mathrm{P}}^{\theta_m(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}d\theta_m,
\end{aligned}$$

where for $j\in\{i,\dots,m\}$,

$$\frac{d\widehat{\mathrm{P}}^{\theta_j(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}=\frac{d\widehat{\mathrm{P}}^{\theta_j(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\big((X_t^{\theta_i^{0+}(\ell)})_{t\in[t_k,t_{k+1}]}\big),$$

and

$$\theta_j(0+):=(\theta_1^0,\dots,\theta_{j-1}^0,\theta_j,\theta_{j+1}^0+\varphi^{j+1}_{n\Delta_n}(\theta^0)u_{j+1},\dots,\theta_m^0+\varphi^m_{n\Delta_n}(\theta^0)u_m).$$

Then, using Girsanov's theorem, we get that

$$
\begin{aligned}
&\widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,x}\left[V\left(\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\left((X_t^{\theta_i^{0+}(\ell)})_{t\in[t_k,t_{k+1}]}\right)-1\right)\right]\\
&=\int_{\theta_i^0+\ell\varphi^i_{n\Delta_n}(\theta^0)u_i}^{\theta_i^0}\widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,x}\Big[V\int_{t_k}^{t_{k+1}}(\partial_{\theta_i}b(\theta_i(0+),t,X_t^{\theta_i^{0+}(\ell)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_i^{0+}(\ell)})\\
&\cdot\left(dB_t-\sigma^{-1}(t,X_t^{\theta_i^{0+}(\ell)})\big(b(\theta_i(0+),t,X_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t,X_t^{\theta_i^{0+}(\ell)})\big)dt\right)\frac{d\widehat{\mathrm{P}}^{\theta_i(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\Big]d\theta_i\\
&+\int_{\theta_{i+1}^0+\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}}^{\theta_{i+1}^0}\widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,x}\Big[V\int_{t_k}^{t_{k+1}}(\partial_{\theta_{i+1}}b(\theta_{i+1}(0+),t,X_t^{\theta_i^{0+}(\ell)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_i^{0+}(\ell)})\\
&\cdot\left(dB_t-\sigma^{-1}(t,X_t^{\theta_i^{0+}(\ell)})\big(b(\theta_{i+1}(0+),t,X_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t,X_t^{\theta_i^{0+}(\ell)})\big)dt\right)\frac{d\widehat{\mathrm{P}}^{\theta_{i+1}(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\Big]d\theta_{i+1}\\
&+\cdots+\int_{\theta_m^0+\varphi^m_{n\Delta_n}(\theta^0)u_m}^{\theta_m^0}\widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,x}\Big[V\int_{t_k}^{t_{k+1}}(\partial_{\theta_m}b(\theta_m(0+),t,X_t^{\theta_i^{0+}(\ell)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_i^{0+}(\ell)})\\
&\cdot\left(dB_t-\sigma^{-1}(t,X_t^{\theta_i^{0+}(\ell)})\big(b(\theta_m(0+),t,X_t^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),t,X_t^{\theta_i^{0+}(\ell)})\big)dt\right)\frac{d\widehat{\mathrm{P}}^{\theta_m(0+)}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\Big]d\theta_m\\
&=\int_{\theta_i^0+\ell\varphi^i_{n\Delta_n}(\theta^0)u_i}^{\theta_i^0}\widehat{\mathrm{E}}^{\theta_i(0+)}_{t_k,x}\Big[V\int_{t_k}^{t_{k+1}}(\partial_{\theta_i}b(\theta_i(0+),t,X_t^{\theta_i(0+)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_i(0+)})dB_t^{\widehat{\mathrm{P}}^{\theta_i(0+)}_{t_k,x}}\Big]d\theta_i\\
&+\int_{\theta_{i+1}^0+\varphi^{i+1}_{n\Delta_n}(\theta^0)u_{i+1}}^{\theta_{i+1}^0}\widehat{\mathrm{E}}^{\theta_{i+1}(0+)}_{t_k,x}\Big[V\int_{t_k}^{t_{k+1}}(\partial_{\theta_{i+1}}b(\theta_{i+1}(0+),t,X_t^{\theta_{i+1}(0+)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_{i+1}(0+)})dB_t^{\widehat{\mathrm{P}}^{\theta_{i+1}(0+)}_{t_k,x}}\Big]d\theta_{i+1}\\
&+\cdots+\int_{\theta_m^0+\varphi^m_{n\Delta_n}(\theta^0)u_m}^{\theta_m^0}\widehat{\mathrm{E}}^{\theta_m(0+)}_{t_k,x}\Big[V\int_{t_k}^{t_{k+1}}(\partial_{\theta_m}b(\theta_m(0+),t,X_t^{\theta_m(0+)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_m(0+)})dB_t^{\widehat{\mathrm{P}}^{\theta_m(0+)}_{t_k,x}}\Big]d\theta_m.
\end{aligned}
$$

Here, for $j\in\{i,\dots,m\}$ the process $B^{\widehat{\mathrm{P}}^{\theta_j(0+)}_{t_k,x}}=(B_t^{\widehat{\mathrm{P}}^{\theta_j(0+)}_{t_k,x}},t\in[t_k,t_{k+1}])$ is a Brownian motion under $\widehat{\mathrm{P}}^{\theta_j(0+)}_{t_k,x}$, where for any $t\in[t_k,t_{k+1}]$,

$$
B_t^{\widehat{\mathrm{P}}^{\theta_j(0+)}_{t_k,x}}:=B_t-\int_{t_k}^t\sigma^{-1}(s,X_s^{\theta_i^{0+}(\ell)})\big(b(\theta_j(0+),s,X_s^{\theta_i^{0+}(\ell)})-b(\theta_i^{0+}(\ell),s,X_s^{\theta_i^{0+}(\ell)})\big)ds.
$$

Next, using Hölder's and Burkholder-David-Gundy's inequalities, conditions **(A2)** and **(A3)**(b), and Lemma 3.4 (ii), we get that

$$
\begin{aligned}
&\left|\widehat{\mathrm{E}}^{\theta_i^{0+}(\ell)}_{t_k,x}\left[V\left(\frac{d\widehat{\mathrm{P}}^{\theta^0}_{t_k,x}}{d\widehat{\mathrm{P}}^{\theta_i^{0+}(\ell)}_{t_k,x}}\left((X_t^{\theta_i^{0+}(\ell)})_{t\in[t_k,t_{k+1}]}\right)-1\right)\right]\right|\\
&\leq\left|\int_{\theta_i^0+\ell\varphi^i_{n\Delta_n}(\theta^0)u_i}^{\theta_i^0}\left|\widehat{\mathrm{E}}^{\theta_i(0+)}_{t_k,x}\Big[V\int_{t_k}^{t_{k+1}}(\partial_{\theta_i}b(\theta_i(0+),t,X_t^{\theta_i(0+)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_i(0+)})dB_t^{\widehat{\mathrm{P}}^{\theta_i(0+)}_{t_k,x}}\Big]\right|d\theta_i\right|
\end{aligned}
$$

$$
\begin{aligned}
&+\left|\int_{\theta_{i+1}^0+\varphi_{n\Delta_n}^{i+1}(\theta^0)u_{i+1}}^{\theta_{i+1}^0}\left|\widehat{\mathrm{E}}_{t_k,x}^{\theta_{i+1}(0+)}\left[V\int_{t_k}^{t_{k+1}}(\partial_{\theta_{i+1}}b(\theta_{i+1}(0+),t,X_t^{\theta_{i+1}(0+)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_{i+1}(0+)})dB_t^{\widehat{\mathrm{P}}_{t_k,x}^{\theta_{i+1}(0+)}}\right]\right|d\theta_{i+1}\right|\\
&+\cdots+\left|\int_{\theta_m^0+\varphi_{n\Delta_n}^{m}(\theta^0)u_m}^{\theta_m^0}\left|\widehat{\mathrm{E}}_{t_k,x}^{\theta_m(0+)}\left[V\int_{t_k}^{t_{k+1}}(\partial_{\theta_m}b(\theta_m(0+),t,X_t^{\theta_m(0+)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_m(0+)})dB_t^{\widehat{\mathrm{P}}_{t_k,x}^{\theta_m(0+)}}\right]\right|d\theta_m\right|\\
&\leq C\Big|\int_{\theta_i^0+\ell\varphi_{n\Delta_n}^i(\theta^0)u_i}^{\theta_i^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_i(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}\\
&\cdot\left(\Delta_n^{\frac{p}{2}-1}\int_{t_k}^{t_{k+1}}\widehat{\mathrm{E}}_{t_k,x}^{\theta_i(0+)}\left[\left|\partial_{\theta_i}b(\theta_i(0+),t,X_t^{\theta_i(0+)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_i(0+)})\right|^p\right]ds\right)^{\frac{1}{p}}d\theta_i\Big|\\
&+C\Big|\int_{\theta_{i+1}^0+\varphi_{n\Delta_n}^{i+1}(\theta^0)u_{i+1}}^{\theta_{i+1}^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_{i+1}(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}\\
&\cdot\left(\Delta_n^{\frac{p}{2}-1}\int_{t_k}^{t_{k+1}}\widehat{\mathrm{E}}_{t_k,x}^{\theta_{i+1}(0+)}\left[\left|\partial_{\theta_{i+1}}b(\theta_{i+1}(0+),t,X_t^{\theta_{i+1}(0+)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_{i+1}(0+)})\right|^p\right]ds\right)^{\frac{1}{p}}d\theta_{i+1}\Big|\\
&+\cdots+C\Big|\int_{\theta_m^0+\varphi_{n\Delta_n}^{m}(\theta^0)u_m}^{\theta_m^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_m(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}\\
&\cdot\left(\Delta_n^{\frac{p}{2}-1}\int_{t_k}^{t_{k+1}}\widehat{\mathrm{E}}_{t_k,x}^{\theta_m(0+)}\left[\left|\partial_{\theta_m}b(\theta_m(0+),t,X_t^{\theta_m(0+)}))^*(\sigma^*)^{-1}(t,X_t^{\theta_m(0+)})\right|^p\right]ds\right)^{\frac{1}{p}}d\theta_m\Big|\\
&\leq C\sqrt{\Delta_n}\,(1+|x|)\left(\left|\int_{\theta_i^0+\ell\varphi_{n\Delta_n}^i(\theta^0)u_i}^{\theta_i^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_i(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}d\theta_i\right|\right.\\
&\qquad+\left|\int_{\theta_{i+1}^0+\varphi_{n\Delta_n}^{i+1}(\theta^0)u_{i+1}}^{\theta_{i+1}^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_{i+1}(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}d\theta_{i+1}\right|\\
&\qquad\left.+\cdots+\left|\int_{\theta_m^0+\varphi_{n\Delta_n}^{m}(\theta^0)u_m}^{\theta_m^0}\left(\widehat{\mathrm{E}}_{t_k,x}^{\theta_m(0+)}\left[|V|^q\right]\right)^{\frac{1}{q}}d\theta_m\right|\right),
\end{aligned}
$$

for some constant $C>0$, where $p,q>1$ and $\frac{1}{p}+\frac{1}{q}=1$. Thus, the result follows. □

References

[1] Basawa, I.V. and Prakasa Rao, B.L.S. (1980), *Statistical Inference for Stochastic Processes*, London: Academic Press.

[2] Barczy, M. and Pap, G. (2010), Asymptotic behavior of maximum likelihood estimator for time inhomogeneous diffusion processes, *Journal of Statistical Planning and Inference*, **140**, 1576-1593.

[3] Barndorff-Nielsen, O. E. and Sørensen, M. (1994), A review of some aspects of asymptotic likelihood theory for stochastic processes, *Int. Stoch. Review*, **62**, 133-165.

[4] Bishwal, J.P.N. (2008), *Parameter Estimation in Stochastic Differential Equations*, Lecture Notes in Mathematics, Springer.

[5] Dietz, H. M. and Kutoyants, Y.A. (2003), Parameter estimation for some non-recurrent solutions of SDE, *Statistics & Decisions*, **21**, 29-45.

[6] Gobet, E. (2001), Local asymptotic mixed normality property for elliptic diffusions: a Malliavin calculus approach, *Bernoulli*, **7**, 899-912.

[7] Gobet, E. (2002), LAN property for ergodic diffusions with discrete observations, *Ann. I. H. Poincaré*, **38**, 711-737.

[8] Hájek, J. (1972), Local asymptotic minimax and admissibility in estimation, *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. I: Theory of statistics*, 175–194.
[9] Höpfner, R. (2014), *Asymptotics statistics*, De Gruyter.
[10] Jacod, J. (2006), Parametric inference for discretely observed non-ergodic diffusions, *Bernoulli*, **12**(3), 383-401.
[11] Jacod, J. (2012), *Statistics and High Frequency Data*, In: M. Kessler, A. Lindner, and M. Sørensen (Eds.), Statistical Methods for Stochastic Differential Equations, Chapman & Hall/CRC Monographs on Statistics and Applied Probability, Volume 124.
[12] Jeganathan, P. (1982), On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal, *Sankhyā Ser. A*, **44**(2), 173–212.
[13] Kutoyants, Y.A. (2004), *Statistical Inference for Ergodic Diffusion Processes*, Springer Series in Statistics, Springer-Verlag, London.
[14] Kunita, H. (1997), *Stochastic Flows and Stochastic Differential Equations. Cambridge Studies in Advanced Mathematics*, Cambridge: Cambridge Univ. Press.
[15] Le Cam, L. (1960), Locally asymptotically normal families of distributions, *Univ. California, Publ. Statist*, **3**, 37-98.
[16] Le Cam, L. and Lo Yang, G. (1990), *Asymptotics in statistics: Some basic concepts*, Springer Series in Statistics. Springer-Verlag, New York.
[17] Luschgy, H. (1992), Local asymptotic mixed normality for semimartingale experiments, *Probab. Theory Related Fields*, **92**(2), 151-176.
[18] Mishra, M. N. and Prakasa Rao, B.L.S. (1985), Asymptotic study of the maximum likelihood estimator for non-homogeneous diffusion processes, *Statistics & Decisions*, **3**, 193-203.
[19] Nualart, D. (2006), *The Malliavin Calculus and Related Topics*, Second Edition, Springer.
[20] Shimizu, Y. (2012), Local asymptotic mixed normality for discretely observed non-recurrent Ornstein-Uhlenbeck processes, *Ann. Inst. Stat. Math*, **64**, 193-211.
[21] Shimizu, Y. (2012), Estimation of parameters for discretely observed diffusion processes with a variety of rates for information, *Ann. Inst. Stat. Math*, **64**, 545-575.
[22] Van Zanten, H. (2000), A multivariate central limit theorem for continuous local martingales, *Statistics & Probability Letters*, **50**(3), 229-235.

Hoang-Long Ngo, Hanoi National University of Education, 136 Xuan Thuy, Cau Giay, Hanoi, Vietnam

E-mail address: ngolong@hnue.edu.vn

Ngoc Khue Tran, Department of Natural Science Education - Pham Van Dong University, 509 Phan Dinh Phung, Quang Ngai City, Quang Ngai, Vietnam

E-mail address: tnkhueprob@gmail.com